CESifo WORKING PAPERS

7597
2019

April 2019

Think global, act local! A mechanism for global commons and mobile firms

Lassi Ahlvik, Matti Liski

CESifo

Impressum:

CESifo Working Papers
ISSN 2364-1428 (electronic version)
Publisher and distributor: Munich Society for the Promotion of Economic Research - CESifo GmbH
The international platform of Ludwigs-Maximilians University's Center for Economic Studies and the ifo Institute
Poschingerstr. 5, 81679 Munich, Germany
Telephone +49 (0)89 2180-2740, Telefax +49 (0)89 2180-17845, email office@cesifo.de
Editor: Clemens Fuest
www.cesifo-group.org/wp

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

CESifo *Working Paper No. 7597*
Category 10: Energy and Climate Economics

Think global, act local! A mechanism for global commons and mobile firms

Abstract

It is tricky to design local regulations on global externalities, especially so if firms are mobile. We show that when costs and outside options are firms' private information, the threat of firm relocation leads to local regulations that are stricter, not looser. This result is general and follows because policy-driven information rents act as targeted compensations to firms that can efficiently limit the externality. The optimal mechanism supplements this strict local regulation with a looser opt-in scheme, creating a global cap for externalities for a subset of firms. We illustrate the magnitude of these effects by providing a quantification of the optimal mechanism for the key sectors in the EU emissions trading system.

JEL-Codes: D820, L510, Q540, Q580.

Keywords: externalities, mechanism design, private information, climate change, emissions trading, carbon leakage.

Lassi Ahlvik
Economics Department
Norwegian School of Economics
Norway - 5045 Bergen
lassi.ahlvik@nhh.no

Matti Liski
Economics Department
Aalto University
Finland - 00076 Aalto
matti.liski@aalto.fi

March 29, 2019
We thank seminar participants at presentations in Athens (EAERE), Bergen, Gothenburg, Helsinki, Paris, Oslo, Torino (IEARE keynote), Toulouse (11th TSE conference on climate and energy), and Zurich for valuable comments and discussion. We also thank Chang-Koo Chi, Baran Doda, Bård Harstad, Daniel Hauser, Torfinn Harding, Johannes Hörner, Yuanhao Li, Jussi Lintunen, Juan-Pablo Montero, Matias Pietola, Julia Salmi, Thomas Sterner, Morten Sæthre, Juuso Välimäki, Po Yin Wong, and Andrey Zhukov for valuable comments and discussions. Matti is grateful for funding from the Academy of Finland program New Energy, and from the Aalto University project SAGA.

1 Introduction

In a single World, externalities such as those related to climate change are truly global but the policies are not. Only about 14% of global carbon emissions are currently subject to some form of carbon pricing, with widely different prices in regional trading systems.[1] Regional climate policies are commonly opposed on the grounds that a local price on the global externality would force businesses out to non-regulated regimes, thereby undermining the effectiveness of the local policies. For example, the U.S. Congress passed a resolution opposing a carbon tax on the basis that, among other things, it *"will lead to more jobs and businesses moving overseas"*.[2] In the European Union, industries have argued that, in the absence of a global climate policy, strengthening the Emissions Trading Scheme would force businesses to leave *"without any environmental need"*.[3]

This paper shows that these arguments do not hold water: A threat of firm relocation calls for stricter local regulations, not looser. We formalize this argument using a mechanism design approach where firms are privately informed about, firstly, the cost of limiting the externality and, secondly, the true costs of moving. Information asymmetries are, in actuality, central to the policy problem as they prevent policy makers from identifying firms for which targeted policy measures can make a difference — the schemes carried out to prevent industry relocation are believed to create large private rents at the expense of public funds.[4]

The optimal mechanism that we characterize strikes a balance between limiting the global externality, avoiding firm relocation, and saving on public funds. A key insight is that a firm's social value in the commons problem depends on its private cost of limiting the externality.

[1]Landmark programs are the ones in California (\$16/tCO_2) and European Union (\$18/tCO_2) but externality prices are also implemented through taxes in individual countries, including UK (\$23/tCO_2), Finland (\$71/tCO_2), France (\$51/tCO_2), and Sweden (\$127/tCO_2). All prices nominal USD as of January 1, 2019. Source: https://carbonpricingdashboard.worldbank.org/map_data.

[2]Source: House Concurrent Resolution 119, 2018. "Expressing the sense of Congress that a carbon tax would be detrimental to the United States economy."

[3]Source: Fagan-Watson, Ben. 2015. "Big business using trade groups to lobby against EU climate policy", Apr 15. https://www.theguardian.com/sustainable-business/2015/apr/15/big-business-trade-groups-lobby-against-eu-climate-change

[4]The total windfall for the European industries from compensation schemes planned to prevent industry relocation has been argued to be as high as 24 billion euros. Source: Krukowska, Ewa. 2016. "EU industry got \$27 billion carbon plan windfall, study says." *Bloomberg*, March 15. https://www.bloomberg.com/news/articles/2016-03-14/eu-industry-got-27-billion-cap-and-trade-windfall-study-says.

Although the policy maker lacks the means to identify the low-cost firms precisely, policies can be used to shape the firm distributions across locations. A lower local price put on the externality benefits the firms that end up paying the price instead of limiting the externality, thereby acting as a support to dirty firms whose location decisions are not relevant for the global commons problem. In the same vein, a higher externality price tends to shape the distribution of firms so that a greater fraction of firms that can efficiently limit the externality end up staying.

Our findings are at odds with the general line of results from the incentive regulation theory (Laffont and Tirole, 1993). There, a downward distortion arises in the regulatory stringency to limit the rents to the low-cost firms and thus to save on public funds spent in the implementation. As in Lewis (1996), this distortion leads to a lower than socially optimal (Pigouvian) price for emissions. In contrast, we show that when firms are not "cornered" to stay, the information rents act as a support to firms that can efficiently limit the externality. This in itself always increases the local externality price, and, under certain conditions, raises it even above the Pigouvian first-best levels. Firms also receive information rents in the other dimension of their private information, their cost of relocation. To limit these rents, our solution to the information problem calls for some firm relocation as an optimal outcome.

For a global externality, it is necessary to *think globally, act locally* but, interestingly, the optimal mechanism allows the policy maker to *act globally.* In principle, the policy maker can eliminate the "leakage" of emissions altogether by incentivizing firms to limit externalities regardless of their location. In fact, if firms' costs were known to the policy maker, the Pigouvian principle would call for a single global externality price for the staying and leaving firms. Without this information, however, the global mechanism leads to novel incentive problems breaking this uniform-price result. The global mechanism optimally manipulates firms' outside options by varying the information rents that firms can expect by moving. The optimal policy is neither purely local or purely global but, as it turns out, it always implements two distinct externality prices: a higher local price for firms that stay, and a lower global price for firms that relocate.

Depending on the industry, it may well be that the abatement and relocation costs are correlated, although it is not *a priori* clear if such a correlation exists and if it is positive or negative. Recent empirical literature on the impact of regulations on industrial activity has noted the importance of such correlation as a confounding factor (see, e.g., Fowlie *et al.*

2016).[5] We find that both the strength and sign of correlation are important for the optimal policy design. Under negative correlation, firms with low abatement costs are the least mobile: The externality prices are pushed down, almost as if firms were immobile. Under positive correlation, on the contrary, low-cost firms are prone to move and, then, it is optimal to elevate the externality price, even above the Pigouvian first-best levels. The global mechanism is particularly valuable if the correlation is positive. Then, moving firms have low cost of limiting the externality and a global price can buy these actions efficiently. Our main result holds in the analysis of perfect correlation: The firm mobility as such can only increase the externality price.

Our theory captures the design problem individual countries are facing if they signed the 2015 Paris Agreement on climate change: Each country (or cooperating region) should unilaterally contribute to the global commons, without a top-down mechanism. When other regions also start contributing, the best-responding unilateral mechanism changes interestingly; mobile and privately informed firms lead to strategic interactions that are new in the literature. If the destination of the mobile firms introduces a fixed externality price, such as a carbon tax, the total externality produced abroad can be manipulated strategically through the distribution of relocating firms. In contrast, when the destination introduces a quantity-based regulation, such as an emissions trading scheme, the total externalities produced abroad are unaffected by firm relocation but the foreign externality price, and thereby firms' outside options, becomes endogenous to the mechanism. We shed some light on these rich interactions, arising from the information economics of the policy problem.

Last, we provide an illustrative quantification of the optimal carbon leakage policy for key sectors in the EU emissions trading system (EU ETS) based on the firm-level data on relocation propensities from Martin *et al.* (2014). The data allow us to draw representative relocation risk distributions for five key sectors that together form 62 per cent of the total industry emissions covered by the trading program. With representative values for the social cost of carbon emissions and the social cost of public funds, we can quantify the optimal carbon leakage, distortions in the emissions price, and the fraction of the sectoral cost that is optimally covered from public funds. The main theoretical results of the paper turn out

[5] Ederington *et al.* (2005) and Naegele and Zaklan (2019) find that the least "footloose" firms are the ones with the largest pollution abatement costs, suggesting a positive correlation between moving and abatement costs. In contrast, in a note written together with Ralf Martin we look at the firm level EU data from Martin *et al.* (2014) and find evidence for a negative correlation (Ahlvik *et al.*, 2017).

to be also economically significant. The optimal local carbon prices are increased upwards by $21 - 31$ per cent compared to the benchmark without firm relocation. The optimal global cap is generally loose, with the global price being always less than a third of the local price. The higher carbon prices also translate into larger cuts, even after the leakage of emissions ($2 - 21\%$ per sector) is taken into account: the threat of relocation, in itself, calls for 3.6 $MtCO_2$ additional emission reductions (15% higher than in the benchmark, an amount roughly equal to total emissions from the manufacturing sector in Denmark), and the optimal global mechanism supplements this by reducing additional 0.5 $MtCO_2$ abroad (2% compared to the benchmark).

Literature. Global commons and mobile firms present a new mechanism design problem because: (i) the principal's welfare depends on the agent's actions even when the agent chooses not to participate in the mechanism; and (ii) the agent's outside option becomes partially endogenous in a global mechanism. The first property is fundamentally different from, say, bankruptcy as relocating firms will continue to produce externalities impacting the principal. The firm's opportunity to move (to face no or other regulation abroad) introduces a novel participation constraint that is different from the zero-profit (Spulber, 1988; Kim and Chang, 1993) or no-loss conditions (Lewis, 1996; Montero, 2000), typically used in the literature on optimal regulation of privately informed firms.[6] Such participation conditions act very similarly in that they only bind for high-cost firms who are left with no information rents. Therefore, the stringency of the regulation (intensive margin) is distorted to limit information rents; high-cost firms can be left with zero rents. In contrast, in our setting with global externalities and privately known relocation costs, policy-driven information rents have social value as such because they act as targeted compensations to those who take actions; in addition, policy-driven relocation (extensive margin) is a tool for limiting the rents of those who don't take actions.[7] Conceptually, the setting leads to a self-selection model with random participation following Rochet and Stole (2002). However, the main results, that is, upward distortion in the regulatory stringency and incentives to non-participating agents, are unique to our externality problem. Similar results do not arise in applications

[6] Zero-profit condition ensures that all firms prefer to stay in the market. No-loss condition means that firms should make no loss relative to the counterfactual situation.

[7] The latter channel of limiting information rents is similar as excluding consumers from using a public good (Hellwig, 2003; Norman, 2004), or preventing natural monopolies from serving a market (Baron and Myerson, 1982).

to firm competition (Rochet and Stole, 2002) or optimal income taxation (Lehmann *et al.*, 2014).

Our study is the first to take a mechanism design approach in the literature on optimal environmental policies under firm relocation, the so-called "carbon leakage" problem. Given that firms' private information on the relocation propensity is an indisputably essential feature of the problem, it is surprising that it has received little attention to date.[8] The earlier literature has restricted the set of policy instruments at one's disposal, for example, to carbon taxes (Markusen *et al.*, 1993; Motta and Thisse, 1994; Hoel, 1997; Ulph and Valentini, 1997; Petrakis and Xepapadeas, 2003; Greaker, 2003). A few studies have focused on limiting firm relocation with lump-sum compensations as the only instrument (Schmidt and Heitzig, 2014; Martin *et al.*, 2014). With such limitations, the policy maker is forced to solve two problems, managing both externalities and relocation, with one instrument so the outcome depends on how exactly the available policy instruments are introduced.[9] By taking a mechanism design approach, we avoid ad-hoc restrictions on the set of admissible policies; rather, the policy maker is left only with constraints stemming from the private information held by firms.

2 The set-up

Consider a continuum of firms with unit mass, each characterized by cost $c \in [\underline{c}, \overline{c}]$ (with $\underline{c} \geq 0$) of reducing one unit of a negative externality, which we refer to as emissions. Each firm's home location is i, and the alternative location is denoted by j. The mass of firms in location i (in location j, resp.) is characterized by density distribution function $\phi_i(C(c), c)$ that depends on firm's type c and also on $C(c) = C_i(c) - C_j(c)$, where $C_k(c)$ is the cost of regulations in location $k = i, j$.[10] The policy maker in i chooses a mechanism, denoted by

[8]Greaker (2003) and Martin *et al.* (2014) note the information problem but leave the mechanism design problem open for future research.

[9]For example, arming the policy maker only with an emission tax means that, in order to compensate the footloose industries, the policy maker must decrease the emission tax (carbon price). But if the policy maker only can only use a subsidy for emission reduction, firms can be compensated by increasing the subsidy level (and thus the effective carbon price). Our general mechanism naturally incorporates both of these policies, among many others.

[10]What we define as a "firm" can be interpreted more broadly as a unit of production such that abatement costs are independently distributed across production units that can be relocated individually. A real-world

$M_i(c)$, implementing, as explicated below, costs and actions for each firm type c. Formally, $M_i(c) = \{C_i(c), X_i(c)\}$, where the required action is $X_i(c) \in [0,1]$, telling how much the firm should limit the externality ($X_i(c) = 1$ for full abatement), and $C_i(c) = cX_i(c) - T_i(c)$ where $T_i(c)$ is the transfer, possibly negative if the firm is taxed. In location j, firms face treatment $M_j(c) = \{C_j(c), X_j(c)\}$, where $M_j(c)$ captures multiple interpretations, with the main ones being: location j is a pollution haven ($X_j(c) = 0$) that may attract firms with subsidies (Section 3); or mechanism $M_j(c)$ might be offered by home location i to attract voluntary participation in j (Section 4).[11]

Local policy maker in i cares about the *local* welfare impacts of *global* emissions, firm's value-added at home, and also the costs of transferring public funds to the firms. The payoff function W_i captures these elements through avoided damages per unit of pollution, $D \geq 0$, firm's location-specific value-added $\gamma \geq 0$, and the cost of public funds $\lambda > 0$:

$$W_i = \int_{\underline{c}}^{\overline{c}} \underbrace{\Big(\gamma + DX_i(c) - C_i(c)\Big)\phi_i(C(c), c)}_{\text{firms staying at location } i} + \underbrace{DX_j(c)\phi_j(C(c), c)}_{\text{firms moving to location } j} - \underbrace{(1+\lambda)T(c)dc}_{\text{cost of funds}}, \qquad (1)$$

where the public funds comprise those spent at home and the transfer used to incentivize firms that move: $T(c) = T_i(c)\phi_i(C(c), c) + T_j(c)\phi_j(C(c), c)$. It is useful to define the *welfare effect of relocation* as the change in social welfare at i when a firm of type c relocates to j:

$$\Delta(M_i(c), M_j(c), c) = -\Big(\gamma + D\big(X_i(c) - X_j(c)\big) - C_i(c) - (1+\lambda)\big(T_i(c) - T_j(c)\big)\Big). \qquad (2)$$

Some insights follow from just observing this definition. First, if a firm cuts emissions in neither regime ($X_i(c) = 0, X_j(c) = 0$), there is no "leakage" of pollution when a firm moves although relocation may still be socially undesirable due to the loss of value-added γ and the firm's possible contribution to public funds. Second, a firm that cuts emissions only when

company can therefore consist of several of these "firms". Our model incorporates both firms actively relocating existing production units to other regimes (analyzed by Martin *et al.* 2014) and investment leakage when regulation causes multinational firms to expand into another location for new production (analyzed by Hanna 2010).

[11] In the main analysis we treat location j as passive with no policies in place. In Section 5.2, we extend to analysis of $M_j(c)$ that is chosen by the policy maker in j to limit the externality. In general, locations i and j can be thought of as individual countries, for example "North" and "South". Alternatively, i could be a climate coalition collectivity maximizing the net welfare of its members, and j a fringe of free-riders. How exactly such a coalition might form is outside the scope of this study; this has been an intensively studied research area since Barrett (1994).

staying ($X_i(c) > 0, X_j(c) = 0$) creates surplus $D - c$, conditional on being in location i. A key observation is that a firm's contribution to the social welfare depends on its privately known cost parameter. Third, when a firm cuts the same in both countries ($X_i(c) = X_j(c) > 0$), relocation has no effect on the externality but the firm's social value still depends on its costs c: All else equal, the local policy maker prefers to keep firms with low costs.

To micro-found the location distribution of firms, $\phi_k(C(c), c)$, $k = i, j$, we assume that firms have two-dimensional private information. That is, in addition to cost of compliance c, each firm has specific privately known relocation cost captured by $\theta \in \mathbb{R}$, so that firms with extreme $\theta > 0$ are cornered to stay in i but some fraction of firms having $\theta < 0$ will move even in the absence of regulation.[12] The density and the cumulative distribution for the relocation costs are $g(\theta)$ and $G(\theta)$, respectively.[13] The distribution of abatement costs follows a continuous density function $f(c)$, with $F(c)$ denoting the respective cumulative distribution. We make the standard regularity requirement on the distributions, stated here for $F(c)$:

Assumption 1. *Distributions $G(\theta)$ and $F(c)$ satisfy the following hazard rate assumption.*[14]

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

A firm with given (θ, c) looks at the cost of compliance under $M_i(c)$ and $M_j(c)$ and decides on which mechanism to report to. In this approach, we follow the techniques developed by Rochet and Stole (2002) and focus on non-stochastic mechanisms: We consider a direct revelation mechanism where firms announce only their cost $\hat{c}$, conditional on reporting to a

[12] The assumption of full support for θ guarantees that there will always be a mass of firms with any given cost c staying in the regime. Relocation costs come from the cost of physically moving and the expected decrease in profits due to choosing another, supposedly less preferable location. The cost of regulation may not be the only reason to relocate production, and there may be other privately observed firm-level shocks making relocation worthwhile even if the cost of regulation was zero; these are captured by negative values of θ. The firm may not know the relocation cost accurately but it is reasonable to assume that the firm has more possibilities to estimate this parameter than the policy maker who, in actuality, may even be uninformed of *where* the firm could potentially move to.

[13] Here, we assume no correlation between relocation and compliance costs $G(\theta|c) = G(\theta)$ but, in Section 5.1, we extend to correlated private information. Another key assumption is that information is purely private, that is, firms have no information about each other which is not available to the policy maker. This prevents the use of mechanisms such as those studied by Cremer and McLean (1988) and Varian (1994).

[14] The condition on the hazard rates is standard (Jullien, 2000) and satisfied for a long list of commonly used distributions (Bagnoli and Bergstrom, 2005).

given mechanism. For any given $M_i(c)$ proposed and taking $M_j(c)$ as given, we then obtain $C(c)$ for each type, and can solve for the mass of firms at location i. In other words, we model firms' relocation decision as an indirect mechanism where no report on relocation cost $\hat{\theta}$ is made.[15]

Cost $C_i(c, \hat{c})$ from reporting $\hat{c}$ for type in mechanism i (respectively in j) depends on the required action and compensation received: $cX_i(\hat{c}) - T_i(\hat{c})$, where the transfer can also be negative. The policy maker thus implements the desired cost level for each type through transfers, taking the cost of funds into account. An incentive-compatible mechanism gives truthful reporting:

$$c = \arg\min_{\hat{c}} \left\{ cX_i(\hat{c}) - T_i(\hat{c}) \right\} \text{ for all } c, \tag{3}$$

defining $C_i(c, c) = C_i(c)$ as the cost of compliance at location i. Thus, effectively, we can write $M_i(c) = \left\{ T_i(c), X_i(c) \right\}$. In addition to the net costs of compliance in each location, the decision to move is also affected by the firm's relocation cost θ. The firm chooses to report in country i if:

$$\theta \geq C_i(c) - C_j(c) \tag{4}$$

The condition means that firms may choose the regime that it is not its most preferred location if the change in costs is high enough to offset the losses from choosing a less-preferred location θ. Equation (3) shows how the policies can be designed to attract firms. On the one hand, a given change in $T_i(c)$ has the same impact across all types of firms. But, on the other hand, a lower $X_i(c)$ attracts firms differently across costs c; namely, the reduction in $X_i(c)$ is more valuable for firms that have *higher* costs. In short, this microstructure gives the mass of firms in locations i and j:

$$\begin{aligned} \phi_i(C(c), c) &= \big(1 - G(C(c))\big) f(c) \\ \phi_j(C(c), c) &= G(C(c)) f(c) \end{aligned} \tag{5}$$

We use $\phi_i'(C(c), c)$ as shorthand for $d\phi_i(C(c), c)/dC(c)$, and define the inverse hazard rate (over C):

$$\eta_i(C(c)) = \frac{\phi_i(C(c), c)}{\phi_i'(C(c), c)}. \tag{6}$$

[15] This is without loss of generality, given the assumption that firms cannot move partially, and the restriction to a deterministic mechanism; see Rochet and Stole (2002). This is in line with real-life regulations that must treat observationally similar firms in identical ways ("same action, same treatment").

Note that $\eta_i(C)$ is negative and increasing in C.

In what follows, we make additional restrictions on mechanism $M_j(c)$; without these, relocation could be stopped by simply setting a large negative $T_j(c)$ (taxes) in location j. In Section 3 we focus on a purely local mechanism by introducing constraint $T_j(c) = 0$ for all c. This can either be seen as a reasonable first step towards the general mechanism, or a political constraint preventing cross-country transfers. This assumption is relaxed in Section 4 where we analyze a general global mechanism and introduce a voluntary participation constraint in location j: $C_j(c) \leq 0$ for all c. After all, as firms move to another sovereign jurisdiction, they cannot be taxed but, still, cross-border rewards for voluntary actions are feasible, which as a real-world instrument is not unheard of.

3 A local mechanism for footloose firms

We begin by analyzing the optimal local mechanism, where the policy maximizes welfare (1) such that equations (3) and (5) hold, and $T_j(c) = 0$ for all c. The last constraint immediately implies that the location j is a "pollution haven", as no reductions are incentivized and thus not made in j, $X_j(c) = 0$.[16] The optimal local policy in i comes down to deciding how much and which firms should limit externalities, and how much of the private cost of regulations are to be covered from the public funds. As we show in the Appendix, the optimal policy can be characterized through two prices: a flat base compensation, T_i^*, for all staying firms, and a top-up for those who cut, $T_i(c) = T_i^* + c_i^*$, where c_i^* is such that, by truthful implementation, only low-cost firms self-select to cut ($X(c) = 1$ for $c \leq c_i^*$) and receive the higher compensation, while the high-cost firms continue producing the externality ($X(c) = 0$ for $c > c_i^*$).

The compensation schedule governs the mass of staying firms at any given c: Changes in $T_i(c)$ translate into changes in the final cost experienced $C_i(c)$ and thus in the moving firm margin $\theta = C_i(c)$. If, as a thought experiment, the policy maker could observe firm types c but not θ, one would select $C_i(c)$ incurred by any given c by choosing a type-tailored base

[16]Even though location j does not limit externalities, it could be active in attracting firms by offering subsidy T_j', not conditional on actions. Then, the relocating firms would be the ones with $\theta \leq C_i(c) + T_j'$, where transfer T_j' is received by the firms that relocate. We can suppress T_j' by interpreting θ' as the net relocation cost $\theta - T_j'$. With this interpretation, the moving firm margin becomes simply $\theta' = C_i(c)$, implying no material change in the analysis.

transfer to optimally save surplus from keeping firms, $\Delta(c)\phi_i(C_i(c), c)$, which would lead to an "elasticity rule",

$$\Delta(c)\phi_i'(C_i(c), c) - \lambda\phi_i(C_i(c), c) = 0 \quad \text{for all } c, \tag{7}$$

where $\Delta(c) = -\big(\gamma + (D_i - c)X_i(c) - \lambda T_i(c)\big)$ is the welfare effect of relocation, as defined in (2). Represented by the first term, a change in $C_i(c)$ impacts relocation, $\phi_i'(C(c), c)$, and thus social loss $\Delta(c)$ arises. As θ is unobservable, cost $C_i(c)$ cannot be changed for a marginal staying firm in isolation, but the same transfer must be paid to all staying firms of the same type c, $-\lambda\phi_i(C(c), c)$.

However, because firms are also privately informed about their costs, compensations cannot be type-tailored for each c but must be given wholesale to all firms taking the same action. In fact, the compensation paid to any firm cutting, $c \leq c_i^*$, depends only on sum $c_i^* + T_i^*$, so a change in cost $C_i(c)$ for such firms can be equivalently implemented through an increase in either c_i^* or T_i^*. The marginal social cost from a change in $C_i(c)$ for all $c \leq c_i^*$ is given by

$$\mu(c_i^*) = \int_{\underline{c}}^{c_i^*} \Big[\Delta(c)\phi_i'(C(c), c) - \lambda\phi_i(C(c), c)\Big] dc. \tag{8}$$

The optimal base transfer, T_i^*, follows from an adjusted elasticity rule that takes the marginal social cost $\mu(c_i^*)$ into account:

$$\Delta(c_i^*)\phi_i'(C(c_i^*), c_i^*) - \lambda\phi_i(C(c_i^*), c_i^*) = -\frac{f(c_i^*)}{1 - F(c_i^*)}\mu(c_i^*). \tag{9}$$

Bearing in mind this reasoning for T_i^*, turn now to the second "price" to be determined, c_i^*. At the margin, reductions generate benefits $D - (1 + \lambda)c_i^*$,[17] for the mass of firms that stay, $\phi_i(C(c_i^*), c_i^*)$. But to incentivize reductions, price c_i^* must be paid to all the firms with $c \leq c_i^*$. The mechanism optimally balances these two effects:

$$\big(D - (1 + \lambda)c_i^*\big)\phi_i(C(c_i^*), c_i^*) + \mu(c_i^*) = 0. \tag{10}$$

The efficient Pigouvian cut-off, $c_i^* = c_P = D/(1+\lambda)$, follows from this condition if $\mu(c_i^*) = 0$. Indeed, if the policy maker could observe c, it would follow the elasticity rule (7) for setting

[17] When private resources are spent on reductions instead of taxes, the actions are costly in terms of public funds, explaining term $1+\lambda$. This term is familiar from the earlier literature, such as Bovenberg and van der Ploeg (1994).

the transfer for each c, and ask reductions for all $c \leq c_i^* = D/(1+\lambda)$. This reduction rule could be indirectly implemented by setting a price on emissions at level $D/(1+\lambda)$ for all staying firms, reflecting the full social cost of emissions. The "price" on emissions, when cost-parameter c is firms' private information, deviates from this Pigouvian reference:

Theorem 1. ***(Local Mechanism)*** *Optimal $M_i(c)$ is characterized by (T_i^*, c_i^*) such that $\{T_i(c) = T_i^*, X_i(c) = 0 : c > c_i^*\}$ and $\{T_i(c) = T_i^* + c_i^*, X_i(c) = 1 : c \leqslant c_i^*\}$. Optimal c_i^* is given by*

$$c_i^* = \frac{D}{1+\lambda} + \Omega_n(c_i^*)$$

where $\Omega_n(c_i^*)$ is a distortion measure for mobile ($n = M$) and immobile firms ($n = I$). If all firms are immobile by assumption, the optimal abatement policy is down-distorted,

$$\Omega_I(c_i^*) = -\frac{\lambda}{1+\lambda}\frac{F(c_i^*)}{f(c_i^*)} < 0.$$

Mobility in itself always increases optimal price c_i^*,

$$\Omega_M(c_i^*) = \frac{\mu(c_i^*)}{\phi_i(C(c_i^*), c_i^*)} > \Omega_I(c_i^*).$$

Proof. See Appendix. □

When firms are immobile, and forced to stay at home location i, the policy maker faces the standard rent extraction - efficiency tradeoff, familiar from incentive regulation (Laffont and Tirole, 1992). This tradeoff manifests itself in Lewis (1996) as a price on emissions that is distorted downwards to limit the information rents of the most efficient firms; rents represent wasted public funds. This is consistent with $\Omega_I(c_i^*) < 0$ in Theorem 1. Footloose firms break this logic, because information rents become an instrument for attracting firms that can efficiently limit the externality. This leads to our main result, that mobility leads to stricter reduction targets. Surprisingly, this result is very general, as it only depends on the standard hazard rate assumptions (Assumption 1) and the fact that public funds are valuable ($\lambda > 0$).[18] Also, not only is a threat of relocation an unsound reason for inaction,

[18] The distortions we find stem from the assumption that public funds are costly ($\lambda > 0$). If they were not ($\lambda = 0$), the policy maker would reach the efficient solution by (i) incentivizing efficient reductions by offering each firm its marginal contribution of their action to the social surplus, following the Vickrey-Clarke-Groves principle as in Dasgupta *et al.* (1980) or Montero (2008), and (ii) keeping all the firms in the regime by offering unlimited compensations.

but the result calls for differentiated externality prices for sectors that are differently exposed to the relocation risk (i.e., sector-specific $G(\theta)$ may differ). In other words, transfers alone cannot deal with differences in the social cost of the relocation risk across sectors.[19]

Building on this logic, the surplus created by the low-cost firms may be so important that it is optimal to set the externality price above the Pigouvian reference, $c_i^* > D/(1+\lambda)$. This happens if in Theorem 1 we have $\Omega_M(c_i^*) > 0$, or equivalently $\mu(c_i^*) > 0$.

Proposition 1. *The policy is stricter than the Pigouvian reference,* $c_i^* > c_P = D/(1+\lambda)$, *if*

$$\underbrace{\lambda\Big(\eta_i(C(c_P)) - \eta_i(C(c))\Big)}_{\substack{\textit{Compliance cost}\\ \textit{effect}}} < \underbrace{D - c - \lambda c_P}_{\substack{\textit{Climate surplus}\\ \textit{effect}}} \qquad \textit{for all } \underline{c} < c \leq c_P. \tag{11}$$

Proof. See Appendix. □

The proposition puts forward two countervailing effects determining the social value of information rents to the low-cost firms. On the one hand, compliance with the regulation is less costly for the low-cost firms so they are less likely to leave at any given level of rents left for them. This is captured by the compliance cost effect. On the other hand, the low-cost firms are socially valuable to keep as they create the largest "climate surplus" by limiting the externality. If this effect dominates the compliance cost effect, the externality price exceeds the Pigouvian level.

To illustrate, recall the elasticity rule, equation (7), which tells us how the policy maker would like to differentiate the treatment of firms for varied c. The transfer given by elasticity rule (7) is illustrated as a dashed line in Figure 1 in two cases. In Figure 1a, the compliance cost effect dominates and the compensation is increasing for all $c < D/(1+\lambda)$. More generous compensations are offered to higher-cost firms, as is typical in incentive regulation. However, Figure 1b shows the opposite case where the climate surplus effect dominates and the type-tailored compensation decreases for all $c < D/(1+\lambda)$. In this latter situation, the policy maker's willingness to compensate firms is inversely related to firms' cost c, because the surplus lost from losing the low-cost firms dominates. Formally, the type-tailored compensation has this property when $\lambda\eta'(C(c)) > 1$ that follows by differentiating condition (11) in Proposition 1.[20]

[19] The heterogeneity of the sectoral leakage risk is a key input to the quantification of Section 6.2.

[20] We characterize the type-tailored compensation schedule formally in the Appendix.

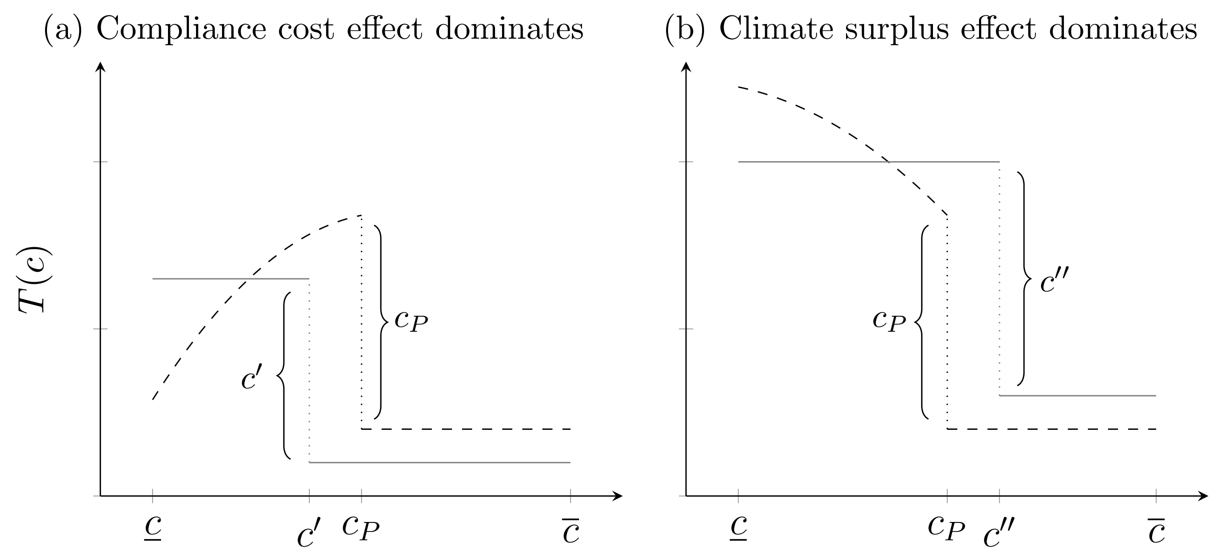

Figure 1: Optimal compensation in the type-tailored (dashed line) case and in the mechanism (solid line), when (a) compliance cost effect dominates, (b) climate surplus effect dominates. The Pigouvian level is denoted by $c_P = D/(1+\lambda)$

Figure 1 depicts the true compensation, under unobservable c, as solid lines. In Figure 1a, the elasticity rule compensation is increasing in $c < c_i^*$ because the policy maker wants to limit rents to the low-cost firms. The optimal contract seeks to achieve this same by distorting the intensive margin, that is, the emissions price downwards ($\mu(c_i^*) < 0$) while also reducing the base compensation to the firms that do not cut. In contrast, if the climate surplus effect dominates, the type-tailored compensation is decreasing in $c < c_i^*$ and the regulator would like to target compensation to the low-cost firms. In the optimal mechanism this is achieved by distorting the intensive margin upwards ($\mu(c_i^*) > 0$) and increasing the base compensation levels.

The appearance of upward distortion is not standard in models of random participation, but it stems from our climate surplus effect, or more exactly, from the fact that firms' privately known type c appears directly in the welfare function through $\Delta(c)$. This is not true in Rochet and Stole (2002) or Lehmann *et al.* (2014); in their settings, assuming a non-decreasing inverse hazard rate (or $\eta'(C(c)) \geq 0$) is sufficient to eliminate any upward distortions.

4 A global mechanism for footloose firms

Until now, we have ruled out cross-bounder transfers and assumed that the policy maker has lost all the possibilities to regulate firms that move. Clearly, this strict focus on purely local policies comes with a loss of generality: After moving, firms' actions continue to impact welfare at home and thus the policy maker has willingness to pay to influence those actions. This improvement in welfare can be achieved by designing one incentive-compatible mechanism for the leaving firms $M_j(c) = \{T_j(c), X_j(c)\}$ and another one for the staying firms $M_i(c) = \{T_i(c), X_i(c)\}$, to maximize the welfare in (1) such that equations (3) and (5) hold and the voluntary participation constraint $C_j(c) \leq 0$ holds for all c.

The global mechanism has characteristics distinct from the purely local one. First, naturally, it is not possible to tax firms located in other sovereign jurisdictions, reflected by the voluntary participation constraint. At home, firms may still be taxed. Second, cross-border transfers have social cost $1 + \lambda$ and thus come with a welfare-loss even if $\lambda = 0$. In contrast, the domestic transfers are effectively evaluated with social cost λ as, intuitively, the transfer circulates within the economy. Last, the outside options of home firms can now be manipulated by the simultaneous offering of the treatments $\{M_i(c), M_j(c)\}$. In particular, a firms' net cost of relocation, $C(c) = c(X_i(c) - X_j(c)) - (T_i(c) - T_j(c))$, becomes independent of the privately known cost c when it cuts in both regimes.

The global mechanism offers, in principle, great opportunities: if $X_i(c) = X_j(c) = 1$ for all low-cost firms $c \leq c^*$, then one implements global price c^* and thereby a global emissions cap. This eliminates the "leakage" problem altogether and the global emissions become independent on firms' location decisions. In fact, if c was observable but not θ, the optimal policy would set the socially optimal global price, $c^* = D/(1 + \lambda)$, both for staying and moving firms, together with differentiated transfers across types and actions. For each $c \leq c^*$ moving, transfer $T_j(c)$ for moving firms would satisfy $C_j(c) = 0$ because there is no need to pay more than necessary to "buy" the socially valuable abatement action from abroad. And, for each $c \leq c^*$ staying, T_i^* would come from the elasticity rule (7).

But when also c is unobservable it is not possible to differentiate transfers $T_j(c)$ to equalize outside options; the rents at home and abroad coexist for any non-marginal firm considering in which location to limit the externality. To introduce the design problem for managing these rents stepwise, consider first policies that cannot discriminate firms' emissions based on their location and is thus constrained to set the same emissions price to all.[21]

Proposition 2. *If the implemented global emission price is constrained to be uniform,* $c^* = c_i^* = c_j^*$*, it is optimally set at:*

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

While Theorem 1 gives the optimal *purely local* mechanism, Proposition 2 is the *purely global* mechanism counterpart. In a sense, the uniform global price corners firms as relocating firms cannot avoid the emission price, almost as if they were immobile. But the outcome is not exactly the same than for immobile firms: only if no firms moved, $G(\cdot) = 0$, externality price c^* would equal the immobile-firm benchmark given in Theorem 1. In contrast, if all firms moved, $G(\cdot) = 1$, c^* would be the optimal monopsony price for buying all reductions with cross-border transfers.

It turns out that the policy maker can always do better than implementing either the purely local (Theorem 1) or the purely global mechanism (Proposition 2). Consider two distinct abatement targets, expressed as local price c_i^* and global price c_j^*, respectively satisfying

$$\big(D - (1+\lambda)c_i^*\big)\phi_i(C(c_i^*), c_i^*) + \mu_i(c_i^*, c_j^*) = 0 \tag{12}$$

$$\big(D - (1+\lambda)c_j^*\big)\phi_j(C(c_j^*), c_j^*) + \mu_j(c_i^*, c_j^*) = 0. \tag{13}$$

These are analogs of (10) for the local mechanism where $\mu(c_i^*)$, the marginal social value of rents left to firms, is now replaced by two marginal valuation terms, one associated with rents caused by the local mechanism and another one with rents caused by the global mechanism:

$$\mu_i(c_i^*, c_j^*) = \int_{\underline{c}}^{c_i^*} \Big[\Delta(c)\phi_i'(C(c), c) - \lambda\phi_i(C(c), c)\Big] dc,$$

$$\mu_j(c_i^*, c_j^*) = \int_{\underline{c}}^{c_j^*} \Big[\Delta(c)\phi_j'(C(c), c) - \lambda\phi_j(C(c), c)\Big] dc,$$

with $\Delta(c)$ from (2). Equation (13) tells that introducing a very small global externality price is welfare improving: it has first-order welfare effects due to reduced emissions, term

[21] As noted before, a real-world company may consist of several of the units that we have called "firms", with headquarters at location i. The policy changes the distribution of activities across locations, and the interpretation of "no discrimination" means that all emissions from the same company is brought under the same cap.

$D-(1+\lambda)c_j^* \approx D$ but only second-order effects to increased relocation and increased public spending, $\mu_j(c_i^*, c_j^*) \approx 0$. Therefore, a small global price is better than the purely local mechanism.

Furthermore, equations (12) and (13) together show that a purely global mechanism would only be optimal if $\mu_i(c_i^*, c_j^*)/\phi_i(C(c_i^*), c_i^*) = \mu_j(c_i^*, c_j^*)/\phi_j(C(c_j^*), c_j^*)$. However, the two expressions are asymmetric by construction: A marginal increase in domestic information rents reduces relocation through $\phi_i'(C(c), c)$, whereas the same increase in information rents at j increases relocation through $\phi_j'(C(c), c)$ by the same amount: $\phi_j'(C(c), c) = -\phi_i'(C(c), c)$, see equation (5). This relocation need not matter for how much the externality is limited but, however, it always matters for the lost the direct benefits (γ) and tax revenues ($-\lambda T_i^*$). To deal with the asymmetry, a global emission cap must therefore balance the reduced emission damages against the aggravated relocation.[22] Formally, we have the following Theorem for the global mechanism:

Theorem 2. ***(Global Mechanism)*** *For global commons, optimal $M_i(c)$ and $M_j(c)$ implement:*

(i) *A strictly positive but downward distorted global price*

$$0 < c_j^* = \frac{D}{1+\lambda} + \frac{\mu_j(c_i^*, c_j^*)}{(1+\lambda)\phi_j(C(c_j^*), c_j^*)} < c^*$$

(ii) *together with a strictly tighter domestic price*

$$c^* < c_i^* = \frac{D}{1+\lambda} + \frac{\mu_i(c_i^*, c_j^*)}{(1+\lambda)\phi_i(C(c_i^*), c_i^*)}$$

where c^* is the uniform-price benchmark defined in Proposition 2.

Under rather minimal assumptions, the optimal mechanism always has both a global ($c_j^* > 0$) and a local ($c_i^* > c_j^*$) component. Our findings emphasize why the policy maker prefers local policies despite the fact that marginal damages are equal across locations. Intuitively, the staying firms are valuable as such, and the local externality price that creates local information rents and thus reduces relocation must be higher than a global price creating rents for firms that move. In fact, Theorem 2 gives a stronger result: the price differentiation leads to a local price that is distorted upwards from the uniform-price benchmark in Proposition 2.

[22] It should be noted that this result also arises without any localized values, $\gamma = 0$, as we show in the Appendix.

5 Extensions

5.1 Correlated private information

If it is known that abatement and moving costs are correlated, then it stands to reason that firms with low costs of limiting the externality can be the most *or least* prone to move, depending on the sign of the correlation. The correlation and its sign, let alone its magnitude, is difficult to reason *a priori*. The current evidence is mixed. Ederington *et al.* (2005) find, using U.S. manufacturing and trade data, that the firms with largest pollution abatement costs also tend to be the least geographically mobile, giving evidence for a positive correlation. Naegele and Zaklan (2019) find evidence for positive correlation using data from the EU Emissions Trading System. In contrast, Levinson and Taylor (2008) find that pollution abatement costs, despite being small fraction of value-added, have an economically significant impact on U.S. trade volume with Canada and Mexico. While not direct evidence, the result is consistent with a negative correlation between mobility and abatement costs.[23]

To introduce correlation in a tractable way, we follow the literature on type-dependent participation constraints (Lewis and Sappington, 1989; Maggi and Rodriguez-Clare, 1995; Jullien, 2000), and consider a perfect correlation between cost c and relocation cost θ:

$$\theta = b + kc. \tag{14}$$

Here b is a scaling constant, and k is publicly known correlation parameter such that for $k > 0$ the lowest cost firms face the lowest relocation cost ($k < 0$ for the reverse). Firms still possess private information but the full correlation renders this information one-dimensional: truthful reporting of the firm's abatement cost immediately reveals also the outside option to the regulator. For ease of exposition, we adopt the convention that the firm's type is c, so that, firms leave when $C_i(c) - C_j(c) > b + kc$.

We now present a set of conclusive results for interior outcomes.[24]

[23] We looked into the question of correlation empirically using the data from Martin *et al.* (2014) in a note written together with Ralf Martin. We find support for a negative correlation between the propensity to move and a measure of abatement costs, contradicting the positive correlation found in Ederington *et al.* (2005).

[24] When the policy maker cares sufficiently about the public funds (λ high enough), it is always optimal to exclude some firms from the regime to save on compensations. Similarly, if externality cost D is sufficiently

Proposition 3. ***(Interior Local Mechanism)*** *Under correlation k, mobile firms continue to face stricter abatement targets than immobile firms (as in Theorem 1):*

(i) *The optimal abatement policy is down-distorted if correlation is negative ($k < 0$),*

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

(ii) *For strong positive correlation ($k > 1$), the policy is distorted above the first-best,*

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

(iii) *For weak positive correlation ($0 < k < 1$), the policy is between situations (i)-(ii).*

A negative correlation means that firms with low abatement costs are "cornered" and, interestingly, the abatement target is pushed down to the immobile-firm benchmark (Lewis, 1996); the policy is the one from Theorem 1 with $\Omega_I(c_i^*)$. But, no matter how strong the correlation is, the policy tightness never goes below this benchmark. In contrast, with positive correlation firms having low abatement cost are the most "footloose", and, then, information rents become a tool for attracting these firm types to stay. This pushes abatement targets up, even above the first-best Pigouvian levels for $k > 1$. To avoid sidetracks we leave the full characterization of the rich economics of the mechanism for the Appendix.[25]

Proposition 4. ***(Interior Global Mechanism)*** *Under correlation k, the global mechanism is implemented only if $k > 0$. For $k > 1$, the global price always emerges:*

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

where domestic c_i^* equals that in Proposition 3, case (ii). For $0 < k < 1$, if in the local mechanism $\gamma < \lambda(T_i^* + c_i^*)$, then global $c_j^* > 0$ is optimal.

large, it will be optimal to require cuts at least from some firms. These two conditions ensure that the mechanism implements an interior outcome. A technical difference to the basic model is that with full correlation, the finite support of c determines the support for θ, by equation (14), and thus the support can no longer be infinite. To facilitate comparison with the results of the basic model, we only focus on interior solutions where some firms cut emissions and some firms move.

[25]For example, weak positive correlation leads to countervailing incentives (Lewis and Sappington, 1989), where firms want to understate their type to emphasize the willingness to leave if it is required to cut the same in both countries ($X_i(c) = X_j(c)$), and overstate their type to emphasize the high compliance costs when only required to cut at home location ($X_i(c) = 1$, $X_j(c) = 0$).

The result echoes the one from the general global mechanism where the global price is also below the domestic one but this result is more specific: negative correlation rules out any benefits from setting a global price. The intuition is straightforward. With $k < 0$, leaving firms have strictly higher costs than the last cutting domestic firm type, c_i^*, so to attract the movers to limit the externality, one would need to set a global price $c_j^* > c_i^*$. Each such reduction unit generates benefit $D-(1+\lambda)c_j^*$ but the cost of buying these reductions exceeds the gains:

$$\big(D-(1+\lambda)c_j^*\big)f(c_j^*)-\lambda F(c_j^*)<0,$$

because for $c_j^* = c_i^*$ the expression on the left is zero by Proposition 3 (case $k < 0$) and thus negative for any $c_j^* > c_i^*$. Even the first global reduction unit would be too costly to buy, given that the domestic mechanism already sets the optimal price. However, with $k > 1$, there is a mass of relocating low-cost firms that will deliver reduction gains $D - (1 + \lambda)c_j^*$ even when the price falls short of the domestic one. The price in Proposition 4 is the optimal uniform monopsony price for procuring reductions from this pool of moving firms. Finally, weak positive correlation ($0 < k < 1$) creates an intricate role for the global price: it allows the policy maker to separate the cutting firms, which could have been kept by sticking to the local mechanism, into two groups. First, very low-cost firms are incentivized to move by offering c_j^*, which saves on implementation costs because $c_j^* < c_i^*$. Second, the policy maker can afford to pay a higher compensation to the staying firms since the information rents are partly capped by the separated low-cost firms. Intuitively, the condition stated limits the social cost of using firms' relocation, γ, relative to the social cost of the transfer, $\lambda(T_i^* + c_i^*)$, to achieve differentiated transfers to cutting firms.

5.2 Policies in location j

So far, the focus has been on how to *unilaterally* implement local policies on global commons in a single country (or a coalition of cooperating countries). But if other countries also start setting policy targets, such as those determined in the Paris Agreement 2015, the world will gradually move towards more global carbon pricing. How does the optimal mechanism in i respond to the increasing policy coverage? The optimal mechanism, it turns out, manipulates the efficacy of policy outcomes elsewhere through the locational distribution of firm types.

This is a source of novel strategic interactions.[26]

To make headway in identifying the strategic channels, we consider $i's$ best responses to two commonly-used policy instruments: a fixed externality price such as a carbon tax, and a fixed externality quantity, such as an emissions trading program.[27] When j adopts a price-based policy instrument, it sets fixed price c_j^* and the total externality reduced in j depends on the number of firms taking action at this price,

$$Q_j(M_i, M_j) = \int_{\underline{c}}^{c_j^*} \phi_j(C(c), c)dc, \tag{15}$$

where we stick to the notation that M_j captures the firm treatment in j although this treatment is now set by the policy maker located in j. Facing regulation in j means that firms have fewer options to escape the regulations as they need to comply in both locations; this effect as such increases the optimal local prices in i. But, as shown in (15), the total level of externality produced in location j becomes endogenous to the mechanism in i: relocating low-cost firms will increase the total reductions delivered by policies in j. This effect tends to decrease to optimal local price in i.

When there is a quantity-based policy in place at j, a fixed quantity Q_j has been set and, then, the externality price $c_j^*(\cdot)$ becomes endogenous to the local mechanism M_i, from:

$$Q_j = \int_{\underline{c}}^{c_j^*(M_i, M_j)} \phi_j(C(c), c)dc. \tag{16}$$

Relocation by low-cost firms reduces the carbon price in j, which in turn decreases the value of firms' outside options and, again, reduces relocation. This effect decreases the optimal externality price implemented by M_i. But, as above with the price instrument, regulation in

[26]Our assumption of constant marginal damage guarantees that there is no traditional strategic common-pool interaction: without leakage, policies are independent between jurisdictions as one region cannot manipulate the *marginal damage* faced by the other through the strategic choice of emissions (see e.g. Van Long 2010). However, firm relocation shifts the *marginal cost* curve, which offers a regime i another way to influence j's emissions. The incentives this strategic channel creates depend on the instrument choice in j. Mideksa and Weitzman (2019) show that the instrument choice creates externalities across jurisdictions under uncertainty. In our set-up, strategic interactions arise even in absence of aggregate uncertainty, creating asymmetry between price- and quantity-based policies even when the marginal cost curves are known.

[27]In principle, countries have flexibility to choose their policy instruments to implement their Nationally Determined Contributions under the Paris Agreement. Some countries (Argentina, South Africa, Singapore) are planning to implement a carbon tax, while others (most importantly China) are planning to launch an emissions trading scheme (Goyal *et al.*, 2018).

j means that it is more difficult for firms to avoid the local emissions reductions, introducing an effect which tends to increase the local externality price in i.

Appendix explicates the above strategic channels and also proves Proposition 5 below. To state the result, define for all $c \leq c_j^* \leq c_i^*$,

$$Z(M_i(c), M_j(c)) = Z(c) = G(C(c))f(c_j^*)D + \Delta(c)g(C(c))F(c_j^*), \tag{17}$$

where $\Delta(c) = -(\gamma - \lambda T_i^* - c - \lambda c_i^*)$ is the net welfare effect of relocation when firms cut emissions in both locations. Expression $Z(c)$ captures two effects through which a marginal increase in c_j^* affects the welfare in i. First, this leads to, all else constant, a mass $G(\cdot)f(\cdot)$ of firms cutting emissions abroad with unit benefit D, the first item in $Z(c)$. Second, a higher foreign price also means that more information rents are being created abroad for low-cost firms, hence a mass of firms $g(\cdot)F(\cdot)$ moves, with $\Delta(c)$ denoting the welfare effect of this relocation.

Proposition 5. *Consider a policy implementing cuts for all $c \leq c_j^* \leq c_i^*$ in location j, either through price- or quantity-based policy. Optimal $M_i(c)$ sets larger c_i^* (smaller, resp.) in response to the price-based policy in j if $Z(c) < 0$ (> 0, resp.) for all c.*

The result captures an interesting difference between price- and quantity-based policies for optimally designed $M_i(c)$. For quantities, the total externality generated by j is fixed, so sending low-cost firms from locality i to comply with $M_j(c)$ crowds out reductions that would have been undertaken by firms in j. Large D tends to make policy c_i^* more ambitious under quantities than prices: elevated information rents keep the low-cost firms at home and thus prevent the crowding out of reductions in j. For prices, the reasoning is reversed since low-cost firms are important in determining the total externality from j.

5.3 General functional forms

The assumptions in our main analysis are simplifying but they turn out not to be very restrictive. First, it would be straightforward to replace our linear damage function with an increasing and convex function $D(X)$ where $X = \int_{\underline{c}}^{\overline{c}} (1-X_i(c))\phi_i(C(c), c) + (1-X_j(c))\phi_j(C(c), c)dc$ gives the total emissions. In this setting, $D'(X)$ naturally replaces the constant marginal damages, denoted by D.[28] Since firms are atomistic, they cannot affect the level of regulation

[28]It should be noted that constant damages can well approximate the predictions of the comprehensive climate-economy models (Golosov *et al.*, 2014; van den Bijgaart *et al.*, 2016).

through the total pollution stock X. From the policy maker's point of view, however, the convex damage function leads to an important difference: Relocation of low-cost facilities shifts the marginal cost curve upwards increasing the optimal emissions price as well as the Pigouvian reference. But since the regulator foresees the aggregate mass of firms that will stay in i given a policy, nothing essentially changes in the regulator's problem.

Next, we have assumed that the economy consists of numerous production units, referred to as "firms", each having a unit cost of emission reductions. As noted earlier, assuming no economies of scale we can interpret "firms" as units of a larger company with independently distributed abatement costs. An alternative modelling approach would be to consider abatement costs that, instead of being independently distributed, depend on a firm-specific privately known technology parameter c so that costs follow a convex function $A(X_k(c), c)$, with $X_k(c) \in \mathbb{R}^+$, $k = i, j$.[29] Then, where $X_i(c)$ and $X_j(c)$ are strictly decreasing, the optimal mechanism sets:[30]

$$\big(D - (1+\lambda)A_x(X_i(c), c)\big)\phi_i(C(c), c) + \mu_i(c) = 0, \tag{18}$$

$$\big(D - (1+\lambda)A_x(X_j(c), c)\big)\phi_j(C(c), c) + \mu_j(c) = 0, \tag{19}$$

where $\mu_i(c) = \int_{\underline{c}}^{c} \Big(\Delta(\tilde{c})\phi_i'(C(\tilde{c}), \tilde{c}) - \lambda\phi_i(C(\tilde{c}), \tilde{c})\Big)d\tilde{c}$ and $\mu_j(c) = \int_{\underline{c}}^{c} \Big(\Delta(\tilde{c})\phi_j'(C(\tilde{c}), \tilde{c}) - (1 + \lambda)\phi_j(C(\tilde{c}), \tilde{c})\Big)d\tilde{c}$. These conditions closely resemble the ones in (12) and (13) in Section 4, with one important difference: while the main mechanism creates two prices, one local and one global, conditions (18) and (19) offer each firms different effective prices depending on their c-type. It follows that, unlike the main mechanisms in Theorems 1 and 2, this cannot be implemented by a simple linear tax or an emissions trading market creating one price for emissions. Although the policy maker can now screen firms better by second-degree price discrimination to save on the public funds, the approach has heavy information requirements that may render it impractical for real-life policy-making: the regulator must be informed about the shape of the abatement cost functions $A(X_k(c), c)$ for each c.

[29] More precisely, we assume $A_x(X, c) > 0$, $A_{xx}(X, c) \geq 0$, $A_{xc}(X, c) > 0$ and that $A(X, c)$ satisfies the Inada conditions.

[30] The abatement levels provided by equations (18) and (19) may sometimes fail to be monotonic, so that the non-monotonicity condition for $X_i(c)$ or $X_j(c)$ is binding and bunching arises: some firms with different types c are offered the same mechanism. A detailed technical analysis of bunching is provided by Rochet and Stole (2002). Following e.g. Lehmann *et al.* (2014), in this extension we focus on the cases where full separation is optimal. Technically, we focus on cases where the non-monotonicity constraints for $X_i(c)$ and $X_j(c)$ never bind.

5.4 Partly local externalities

Climate emissions are undoubtedly global externalities, but their reduction is typically associated with other jointly produced local pollutants, as emphasized by recent empirical work (e.g., Wagner and De Preux 2016; Holland *et al.* 2018). We can include this effect by letting damages from location j be a fraction of the local damages αD, where $\alpha \in [0,1]$ denotes the share of damages that are global. This allows us to write the net loss from relocation in equation (2) as:

$$\Delta(M_i(c), M_j(c), c) = -\Big(\gamma - (1-\alpha)D + D\big(X_i(c) - \alpha X_j(c)\big) - C_i(c) - (1+\lambda)\big(T_i(c) - T_j(c)\big)\Big).$$

Introducing $\alpha < 1$ has two effects. First, as one would expect, the loss of moving is smaller for a more localized damage, captured by smaller α: As firms move, local pollutants move with them, making relocation less harmful for the home regime i even without active policies in j. This is represented by the second term $(1-\alpha)D$. The main results are qualitatively robust to this extensions. Technically, more local damages enter as a fixed, type-independent, damages from moving: We could re-define $\gamma' = \gamma - (1-\alpha)D$, allowing us to follow the derivation of Theorem 1.[31] A key observation is that the net welfare effect of relocation, shown in the equation above, still depends on the firms' cost type c through $C_i(c)$ and, given that some firms move, the regulator would rather keep the clean firms with low c than dirty firms with high c. As a second effect of local damages, a lower α gives less reasons to incentivize the moving firms to reduce their emissions abroad. The local regulator benefits $\alpha D X_j(c)$ when emissions are reduced in the other location: A local component in damages reduces the importance of the global mechanism.

6 Application to the EU ETS sectors

6.1 Practical implementation strategies

We look at the magnitude of the results by providing a quantification of the optimal mechanism for the key sectors in the EU emissions trading system. To this end, putting numbers

[31] As a difference to our main mechanism, for a high enough value of α we may have $\gamma' \leq 0$: a firm's social contribution becomes negative. This resembles the "not in my backyard" result of strategic environmental policy literature with local pollutants where the regulator may have an incentive to drive away some firms to avoid the damages caused by their local pollution (Hoel, 1997; Greaker, 2003).

first aside, it is useful to spell out how the theory mechanism can be mapped into policy instruments with a practical meaning. To mimic the theory outcomes, the practical instruments should implement, first, differentiated base compensations and, second, differentiated effective externality prices across sectors. The base compensation depends on the policy maker's assessment of the sector-specific relocation and abatement cost distributions $F(c)$ and $G(\theta)$, as well as the value created per emissions, γ; more compensation paid to industries under *higher* relocation risk. The base compensation can take the form of direct monetary compensation, free allowances, or lump-sum rebates.[32] The quantification thus needs to specify sector-specific representations for these key drivers of compensations.

To differentiate the externality price across sectors, compensations can be made sensitive to firms' actions and differently so in varied sectors: the price on emissions is given by the compensation lost by not reducing emissions.[33] Compliance cost rebates, or provisions to use of cheaper international or domestic offsets for compliance are examples of such compensations.[34] Lower externality prices should, everything else constant, be given to sectors with a *lower* relocation risk. In particular, our results imply that a threat of relocation is never a valid reason to exempt entire sectors, such as those Emissions Intensive and Trade Exposed (EITE) industries, from regulation. In our model, this would resemble the extreme case where the carbon price, as well as the lump-sum compensation, is set to zero. This is never an optimal policy but, instead, a sector susceptible to leakage should be included in

[32]For example, in the EU total 43 % of the allowances are given away for free during 2013-2020. Sectors deemed to be exposed to a significant risk of carbon leakage receive 100 % of their estimated allowance need for free, whereas the free allocation to non-leakage sector is gradually reduced to 30 % by year 2020. In addition, the most energy-intensive sectors can be given monetary compensation through national state aid schemes (EC, 2019).

[33]To clarify potential sources of confusion: The local mechanism (Theorem 1) can be implemented simply by a cap-and-trade scheme or by a carbon tax – as there is no aggregate uncertainty, these two policy instruments lead to identical outcomes. The differentiation in the effective prices is achieved by transfers conditional on what firms do in the program, that is, if they cut or comply by paying for their emissions.

[34]Rebates from environmental taxes are a common tool for subsidizing energy-intensive industries. For example, in Finland, firms with energy tax bill exceeding .5% of the value-added are entitled to participate in a rebate program. This together with the rebate rule targets the largest energy-consuming firms, see Tamminen *et al.* (2016). Firms in the EU ETS can use international credits generated through Clean Development Mechanism (CDM) and Joint Implementation, up to a polluter-specific percentage, for compliance (EC, 2019). Historically, prices in the CDM markets have been 72% of the allowance price in the EU ETS between 2008-2012 and only 3% between 2013-2016 (https://www.theice.com/index). As offsets are valueless to low-cost firms, allowing their use becomes a targeted compensation to high-cost firms.

regulation and offered adequate lump-sum compensations.[35]

The global mechanism (Theorem 2), in turn, can be implemented by cross-border permit trading, by allowing relocation firms to sell permits to the local policy maker for verified emission reductions in their new location. This is against the main principles in the EU emissions trading system, where moving firms are not permitted to continue trading with the market. In our optimal mechanism, these cross-border trades must be made voluntary and they resemble an "opt-in" scheme where relocating firms with high abatement cost choose not to participate, receiving their outside option. While cross-border transfers may politically difficult to implement, they are not new in the international arena; see for example the Clean Development Mechanism (CDM). Yet, in line with our Theorem 2, this global price is always below the local carbon price.

6.2 Quantification

To illustrate the economic significance of the results, we carry out an exploratory quantification for the key sectors under relocation risk in the EU emissions trading system: cement, iron and steel, chemical and plastic, wood and paper, and glass. Together, these five sectors produce 355 MtCO_2, or 62 per cent of emissions from all industrial installations covered by the EU ETS (EEA, 2017). Our analysis builds on the survey data collected by Martin *et al.* (2014). The data contains firm-level assessments of the relocation probability conditional on receiving no free permits and receiving 80% for free.[36] From these responses, we construct emissions-weighted industry averages for the relocation probability, see Table 1. We choose social cost carbon to be $D = 25$ euros/tCO_2[37], and fit normal distributions for relocation costs, one for each industry, based on the responses. For instance, for "cement", we calibrate the two parameters of the normal distribution using the two relocation probabilities from

[35] This was the case for aviation sector, which was exempt from emissions pricing until 2012, but has since been included although it is receiving most of its allowances for free (EC, 2019).

[36] In the survey, the firms were asked: "Do you expect that government efforts to put a price on carbon emissions will force you to outsource parts of the production of this business site in the foreseeable future, or to close down completely?" and "How would your answer to the previous questions change, if you received a free allowance for 80% of your current emissions?" Answers were given in a Likert scale between 1 and 5, where 1 was no impact (1 %), 3 was significant reduction in production (10 %) and 5 was complete close-down (99 %).

[37] The social cost of carbon 25 €/tCO_2 comes close to the median value in a distribution from integrated-assessment model outputs of 232 distinct studies (van den Bijgaart *et al.*, 2016).

Table 1: Descriptive statistics of the data used

	Total emissions in 2015 ($MtCO_2$)[1]	EBIT per emissions (€/tCO_2)[2]	Relocation probability[2] 0% compen-sation	80% compen-sation	No. firms[2]
Cement	113.8	32.73	0.46	0.20	46
Iron and Steel	120.6	80.52	0.60	0.21	25
Chemical and Plastic	74.9	177.96	0.24	0.06	64
Wood and Paper	27.1	89.31	0.14	0.03	61
Glass	18.2	120.56	0.14	0.05	24

[1]Data from EEA (2017), [2]Data from Martin *et al.* (2014).

Table 1: 46% percent of firms relocate if the full social cost is imposed, and 20% relocate if 80% of the true social cost D that they inflict is actually given back to firms.

Our estimate of parameter γ, the industry-specific value of a firm staying, is based on emissions-weighted average earnings before investment and tax (EBIT) per unit of pollution, expressed as $€/tCO_2$ in Table 1. Abatement cost estimates are hard to come by at the industry level. We use the marginal abatement cost estimates for the EU energy intensive industries from Böhringer *et al.* (2014), which is approximately linear and can be approximated by a uniform distribution.[38] The social cost of public funds is $\lambda = .6$.[39] With these assumptions, computing the optimal mechanism is a straightforward numerical exercise.

We report the optimal policies per sector in Table 2. Panel A gives the optimal local mechanism (Theorem 1). The first column gives the base compensation level, and the second column presents the effective local emissions price per sector. The key sectors are treated very differently. At one extreme, Iron and Steel polluters receive a compensation, whereas the compensation to Wood and Paper is negative; that is, even firms who cut emissions face a net tax (base compensation - local price = -2.2 euros/tCO_2). To further interpret the results, consider the optimal emissions price in the absence of leakage: Using Theorem 1 (immobile firms) we obtain 11.4 euros/tCO_2 for all sectors. The impact of leakage is quantitatively significant: the effective CO_2 price is substantially elevated, by $21 - 31$ per cent compared

[38]From Böhringer *et al.* (2014), we obtain $100MtCO_2$ of reductions at 47 euros/tCO_2 which pins down the slope of the marginal cost curve. Abatement is allocated to sectors in proportion to their unrestricted emissions.

[39]Country-specific circumstances have a large impact on the real costs of taxation so one number cannot fit the entire EU. The chosen number is higher than those presented in Bovenberg and Goulder (1996) but closer to the survey of more recent estimates in Holtsmark and Bjertnæs (2015).

Table 2: Optimal mechanism for the EU ETS sectors

	Implementation of the mechanism			Implied emission reductions		
	Base compensation (€/tCO_2)	Local CO_2 price (€/tCO_2)	Global CO_2 price (€/tCO_2)	Local reductions ($MtCO_2$)	Global reductions ($MtCO_2$)	Emission leakage ($MtCO_2$)
Panel A - Local mechanism						
Cement	-9.6	14.9	-	8.41	-	1.74
Iron and Steel	6.9	13.7	-	9.53	-	0.39
Chemical and Plastic	0.3	13.7	-	6.09	-	0.12
Wood and Paper	-16.5	14.1	-	2.18	-	0.11
Glass	-15.0	14.5	-	1.48	-	0.10
Total				27.69		2.46
Panel B - Global mechanism						
Cement	-9.9	14.6	4.8	8.17	0.48	1.32
Iron and Steel	6.9	13.7	2.1	9.50	0.02	0.37
Chemical and Plastic	0.3	13.7	1.1	6.09	0.01	0.11
Wood and Paper	-16.5	14.1	2.2	2.17	0.01	0.10
Glass	-15.0	14.5	2.5	1.48	0.01	0.09
Total				27.41	0.53	1.99

Notes: Optimal base compensations, implied marginal carbon taxes (panels 1-3) and the effects on emission reductions and leakage (panels 4-6) for the local mechanism (Panel A) and the global mechanism (Panel B). The social cost of carbon is 25 €/tCO_2 and the social cost of public funds is $\lambda = .6$. Assumptions detailed in the text. The first-best (complete information) carbon price is 15.6 €/tCO_2. The immobile-firm benchmark is 11.4 €/tCO_2, leading to 24.12 $MtCO_2$ reductions when no firms relocate.

to the benchmark level with where leakage was assumed away (11.4 euros/tCO_2). Yet, for all the sectors the emissions price falls short of the Pigouvian benchmark (15.6 euros/tCO_2), suggesting that the abatement cost effect dominates for all the sectors (Proposition 1).

In columns 4-6 we show that these higher local prices translate into larger global emission reductions even when firm relocation is taken into account. For a benchmark, if all the sectors considered would be immobile by assumption, the total emission reductions would be 24.12 $MtCO_2$. Under the optimal mechanism, the total reduction 27.69 $MtCO_2$, 15 per cent larger than the benchmark, and the optimal emission leakage is 2.46 $MtCO_2$, or 9 per cent of the total emission reduction. Finally, Panel B presents the optimal global mechanism per sector (Theorem 2). The global price is well below the local one: only $8-33$ per cent of the local price. This global emission cap is leads to additional reductions of 0.53 $MtCO_2$ for a total of 2.4 million euros used in cross-border transfers. In this quantification, the local treatment of firms does not change much if the global price is introduced: the base compensation, local price and the local reductions are relatively robust to the introduction of the global mechanism.

7 Conclusions

Putting a price on global externalities is the economists' approach to the global commons problem – hundred years after the first proposal for corrective prices, the profession continues to believe in the approach (Cramton *et al.*, 2017). We found that these principles should not be given up, though the global action may be lacking behind. In contrast, corrective prices, when suitably designed, attract firms that can contribute to the commons problem to stay. This elevates the optimal corrective price above the level that would be chosen if the policy maker overlooked that not only externalities but also firms are global. Along the same lines of reasoning, self-interested decisions justify payments that, effectively, implement externality prices also for moving firms.

These results advise against regulatory rollbacks and other forms of routinely-used compensation policies that effectively curb carbon prices, including: emission tax refunds, the use of cheap offsets, and exemptions of certain industries from regulation. As firm relocation serves to limit overcompensation paid to industries, observing carbon leakage is not a sign of a failed policy but an essential feature of the information-constrained optimal mechanism. On the contrary, one can argue that the EU emissions trading system has failed exactly because no relocation is observed; see studies by Dechezleprêtre *et al.* (2019) and Naegele and Zaklan (2019). Finally, instead of the current practice where moving firms stop being part of the regulation, there is a well-founded justification for cross-border transfer trading that, effectively, allows moving firms to sell their emission reductions to the local policy maker.

References

Ahlvik, L., Liski, M. and Martin, R. (2017). The association between the vulnerability score and compliance costs. *Mimeo*.

Bagnoli, M. and Bergstrom, T. (2005). Log-concave probability and its applications. *Economic Theory*, **26** (2), 445–469.

BARON, D. P. and MYERSON, R. B. (1982). Regulating a monopolist with unknown costs. *Econometrica*, **50** (4), 911–930.

BARRETT, S. (1994). Self-enforcing international environmental agreements. *Oxford Economic Papers*, pp. 878–894.

BÖHRINGER, C., DIJKSTRA, B. and ROSENDAHL, K. E. (2014). Sectoral and regional expansion of emissions trading. *Resource and Energy Economics*, **37**, 201–225.

BOVENBERG, A. L. and GOULDER, L. H. (1996). Optimal environmental taxation in the presence of other taxes: general-equilibrium analyses. *The American Economic Review*, **86** (4), 985–1000.

— and VAN DER PLOEG, F. (1994). Environmental policy, public finance and the labour market in a second-best world. *Journal of Public Economics*, **55** (3), 349–390.

CRAMTON, P., MACKAY, D. J., OCKENFELS, A. and STOFT, S. (2017). *Global carbon pricing: the path to climate cooperation*. The MIT Press.

CREMER, J. and MCLEAN, R. P. (1988). Full extraction of the surplus in bayesian and dominant strategy auctions. *Econometrica*, pp. 1247–1257.

DASGUPTA, P., HAMMOND, P. and MASKIN, E. (1980). On imperfect information and optimal pollution control. *The Review of Economic Studies*, **47** (5), 857–860.

DECHEZLEPRÊTRE, A., GENNAIOLI, C., MARTIN, R., MUÛLS, M. and STOERK, T. (2019). Searching for carbon leaks in multinational companies. *Centre for Climate Change Economics and Policy Working Paper No. 187 ISSN 2515-5709*.

EC (2019). EU Emissions Trading System (EU ETS). https://ec.europa.eu/clima/policies/ets_en, online; accessed 2019-03-01.

EDERINGTON, J., LEVINSON, A. and MINIER, J. (2005). Footloose and pollution-free. *The Review of Economics and Statistics*, **87** (1), 92–99.

EEA (2017). EU emissions trading system (ETS) data viewer. https://www.eea.europa.eu/data-and-maps/dashboards/emissions-trading-viewer-1.

Fowlie, M. L., Reguant, M. and Ryan, S. (2016). Measuring leakage risk. In *Technical Report, California Air Resources Board*.

Golosov, M., Hassler, J., Krusell, P. and Tsyvinski, A. (2014). Optimal taxes on fossil fuel in general equilibrium. *Econometrica*, **82** (1), 41–88.

Goyal, R., Gray, S., Kallhauge, A. C., Nierop, S., Berg, T. and Leuschner, P. (2018). State and trends of carbon pricing 2018. *The World Bank: Washington, DC, USA*.

Greaker, M. (2003). Strategic environmental policy when the governments are threatened by relocation. *Resource and Energy Economics*, **25** (2), 141–154.

Hanna, R. (2010). US environmental regulation and FDI: evidence from a panel of US-based multinational firms. *American Economic Journal: Applied Economics*, **2** (3), 158–89.

Hellwig, M. F. (2003). Public-good provision with many participants. *The Review of Economic Studies*, **70** (3), 589–614.

Hoel, M. (1997). Environmental policy with endogenous plant locations. *The Scandinavian Journal of Economics*, **99** (2), 241–259.

Holland, S. P., Mansur, E. T., Muller, N. and Yates, A. J. (2018). Decompositions and policy consequences of an extraordinary decline in air pollution from electricity generation. *NBER Working Paper 25339*.

Holtsmark, B. and Bjertnæs, G. H. (2015). The size of the marginal cost of public funds – a discussion with speciel relevance to Norway. *Statistics Norway, Reports 2015/13*.

Jullien, B. (2000). Participation constraints in adverse selection models. *Journal of Economic Theory*, **93** (1), 1–47.

Kim, J.-C. and Chang, K.-B. (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 Tirole, J. (1993). *A theory of incentives in procurement and regulation*. MIT press.

Lehmann, E., Simula, L. and Trannoy, A. (2014). Tax me if you can! Optimal non-linear income tax between competing governments. *The Quarterly Journal of Economics*, **129** (4), 1995–2030.

Levinson, A. and Taylor, M. S. (2008). Unmasking the pollution haven effect. *International Economic Review*, **49** (1), 223–254.

Lewis, T. R. (1996). Protecting the environment when costs and benefits are privately known. *The RAND Journal of Economics*, pp. 819–847.

— and Sappington, D. E. (1989). Inflexible rules in incentive problems. *The American Economic Review*, **27** (4), 69–84.

Maggi, G. and Rodriguez-Clare, A. (1995). On countervailing incentives. *Journal of Economic Theory*, **66** (1), 238–263.

Markusen, J. R., Morey, E. R. and Olewiler, N. D. (1993). Environmental policy when market structure and plant locations are endogenous. *Journal of Environmental Economics and Management*, **24** (1), 69–86.

Martin, R., Muûls, M., De Preux, L. B. and Wagner, U. (2014). Industry compensation under relocation risk: A firm-level analysis of the EU emissions trading scheme. *The American Economic Review*, **104** (8), 2482–2508.

Mideksa, T. and Weitzman, M. L. (2019). Prices vs. quantities across jurisdictions. *Journal of the Association of Environmental and Resource Economists*, p. Forthcoming.

Montero, J.-P. (2000). Optimal design of a phase-in emissions trading program. *Journal of Public Economics*, **75** (2), 273–291.

— (2008). A simple auction mechanism for the optimal allocation of the commons. *The American Economic Review*, **98** (1), 496–518.

Motta, M. and Thisse, J.-F. (1994). Does environmental dumping lead to delocation? *European Economic Review*, **38** (3-4), 563–576.

Naegele, H. and Zaklan, A. (2019). Does the EU ETS cause carbon leakage in european manufacturing? *Journal of Environmental Economics and Management*, **93**, 125–147.

NORMAN, P. (2004). Efficient mechanisms for public goods with use exclusions. *The Review of Economic Studies*, **71** (4), 1163–1188.

PETRAKIS, E. and XEPAPADEAS, A. (2003). Location decisions of a polluting firm and the time consistency of environmental policy. *Resource and energy Economics*, **25** (2), 197–214.

ROCHET, J.-C. and STOLE, L. A. (2002). Nonlinear pricing with random participation. *The Review of Economic Studies*, **69** (1), 277–311.

SCHMIDT, R. C. and HEITZIG, J. (2014). Carbon leakage: Grandfathering as an incentive device to avert firm relocation. *Journal of Environmental Economics and Management*, **67** (2), 209–223.

SEIERSTAD, A. and SYDSÆTER, K. (1987). *Optimal control theory with economic applications*. Elsevier North-Holland, Inc.

SPULBER, D. F. (1988). Optimal environmental regulation under asymmetric information. *Journal of Public Economics*, **35** (2), 163–181.

TAMMINEN, S., OLLIKKA, K. and LAUKKANEN, M. (2016). Suomen energiaverotus suosii energiaintensiivisiä suuryrityksiä. *VATT Policy Brief*, **2-2016**.

ULPH, A. and VALENTINI, L. (1997). Plant location and strategic environmental policy with inter-sectoral linkages. *Resource and Energy Economics*, **19** (4), 363–383.

VAN DEN BIJGAART, I., GERLAGH, R. and LISKI, M. (2016). A simple formula for the social cost of carbon. *Journal of Environmental Economics and Management*, **77**, 75–94.

VAN LONG, N. (2010). *A survey of dynamic games in economics*, vol. 1. World Scientific.

VARIAN, H. R. (1994). A solution to the problem of externalities when agents are well-informed. *The American Economic Review*, pp. 1278–1293.

WAGNER, U. J. and DE PREUX, L. (2016). The co-benefits of climate policy: Evidence from the EU Emissions Trading Scheme. *Beitrge zur Jahrestagung des Vereins fr Socialpolitik 2016: Demographischer Wandel - Session: International Climate Policy, No. D15-V2, ZBW - Deutsche Zentralbibliothek fr Wirtschaftswissenschaften, Leibniz-Informationszentrum Wirtschaft, Kiel und Hamburg.*

APPENDIX FOR ONLINE PUBLICATION

A Appendix: Proofs

Proof of Theorem 1. (Local mechanism)

We begin by introducing a series of lemmas characterizing the optimal mechanism.

Lemma 1. *In a given mechanism i, the transfer $T_i(c)$ is constant when the policy $X_i(c)$ is constant.*

Proof. Proof by contradiction. Assume that there are c and c' with $T_i(c') > T_i(c)$ with $X_i(c') = X_i(c)$. Now firm c can get a lower net cost by reporting c':

$$cX_i(c') - T_i(c') < cX_i(c) - T_i(c)$$

However, this is in violation of the incentive compatibility condition in equation (3). Q.E.D.

Lemma 2. *In a given mechanism i, $X_i(c)$ is nonincreasing in c.*

Proof. Proof by contradiction. If this is not true, there are types c and c', with $c < c'$ and $X_i(c') > X_i(c)$. Incentive compatibility requires $T_i(c)$ and $T_i(c')$ such that types do not want to report the other type:

$$cX_i(c) - T_i(c) \leq cX_i(c') - T_i(c')$$

$$c'X_i(c) - T_i(c) \geq c'X_i(c') - T_i(c')$$

Combining these two inequalities leads to:

$$c'\big(X_i(c') - X_i(c)\big) - T_i(c') \leq c\big(X_i(c') - X_i(c)\big) - T_i(c') \Rightarrow c' \leq c$$

But this is a contradiction. Q.E.D.

Lemma 3. *In the optimal local mechanism ($T_j(c) = 0$, $X_j(c) = 0$), actions take a bang-bang form: $X_i(c) = \{0, 1\}$*

Proof. The objective function (1) can be written as:

$$\max_{X_i(c), C_i(c)} \int_{\underline{c}}^{\overline{c}} \Big(\gamma + DX_i(c) - (1+\lambda)cX_i(c) + \lambda C_i(c)\Big)\phi_i\big(C_i(c), c\big)dc$$

s.t. $C_i'(c) = -X_i(c)$ holds for all c. Denoting the co-state variable by $\mu_i(c)$, the Hamiltonian for this problem reads:

$$\mathcal{H} = \big(\gamma + DX_i(c) - (1+\lambda)cX_i(c) + \lambda C_i(c)\big)\phi_i(C_i(c), c) - \mu_i(c)X_i(c)$$

The Hamiltonian is linear in the controls $X_i(c)$, and the necessary conditions for optimality imply that $X_i(c)$ takes a bang-bang form: $X_i(c) = \{0, 1\}$. Q.E.D.

There exists a solution to the problem as stated in Lemma 3 by Filippov-Cesari Theorem (Theorem 8, page 132, Seierstad and Sydsæter 1987). We find next a solution that satisfies the necessary conditions; it follows by Assumption 1 that there is a unique solution that satisfies the conditions. Lemmas 1-3 above tell us that the optimal policy takes a threshold form, where $X_i(c) = 1$ for $c \leq c_i^*$ and $X_i(c) = 0$ for $c > c_i^*$. Transfers are either $T_i^1(c) = T_i^*$ for $c > c_i^*$ or $T_i^2(c) = T_i^* + c_i^*$ for $c \leq c_i^*$, guaranteeing indifference for type c^*: $-T_i^1(c^*) = c^* - T_i^2(c^*)$. In this formulation, the regulator in country i is left to find c_i^* and T_i^* that maximize the social welfare, given by equation (1):

$$\max_{c_i^*, T_i^*} W_i = \int_{\underline{c}}^{c_i^*} \Big(\gamma + D - c - \lambda(T_i^* + c_i^*)\Big)\phi_i(c - T_i^* - c_i^*, c) + \int_{c_i^*}^{\overline{c}} \Big(\gamma - \lambda T_i^*\Big)\phi_i(-T_i^*, c)dc$$

Here the first integral covers all the firms below the threshold c_i^* cutting emissions, and the second term covers firms above the threshold that do not cut emissions. Begin by taking the first-order condition with respect to c_i^*. Using Leibniz's integral rule, we get:

$$\Big(D - c_i^* - \lambda c_i^*\Big)\phi_i(-T_i^*, c_i^*) + \int_{\underline{c}}^{c_i^*} \Big(\Delta(c)\phi_i'(c - T_i^* - c_i^*, c) - \lambda\phi_i(c - T_i^* - c_i^*)\Big)dc = 0 \qquad \text{(A.1)}$$

where $-\Delta(c) = \gamma + (D - c)X_i(c) - \lambda T_i(c)$ denotes the net welfare effect of relocation by type c. Simplify and solve for c_i^* to derive equation (10):

$$c_i^* = \frac{D}{1+\lambda} + \frac{\mu(c_i^*)}{(1+\lambda)\phi_i(-T_i^*, c_i^*)} \qquad \text{(A.2)}$$

where $\mu(c_i^*)$ is the integral term

$$\mu(c_i^*) = \int_{\underline{c}}^{c_i^*} \Big(\Delta(c)\phi_i'(C(c), c) - \lambda\phi_i(C(c), c)\Big)dc. \qquad \text{(A.3)}$$

Then, find the first-order condition with respect to T_i^*:

$$\underbrace{\int_{\underline{c}}^{c_i^*} \Big(\Delta(c)\phi_i'(c - T_i^* + c_i^*, c) - \lambda\phi_i(c - T_i^* + c_i^*, c)\Big)dc}_{=\mu(c_i^*)} + \int_{c_i^*}^{\overline{c}} \Big(\Delta(c)\phi_i'(-T_i^*, c) - \lambda\phi_i(-T_i^*, c)\Big)dc = 0 \qquad \text{(A.4)}$$

For $c > c_i^*$, $X_i(c) = 0$ and therefore the second integral does not depend on c apart from the term

$f(c)$, see eq. (5), so that (A.4) can be written as:

$$\mu(c_i^*) + \Delta(c_i^*) \underbrace{g(-T_i^*)(-1)}_{=\phi_i'(-T_i^*, c_i^*)/f(c_i^*)} (1 - F(c_i^*)) - \lambda\Big(\underbrace{(1 - G(-T_i^*))}_{=\phi_i(-T_i^*, c_i^*)/f(c_i^*)} (1 - F(c_i^*)\Big) = 0 \tag{A.5}$$

$$\Rightarrow$$

$$\Delta(c_i^*)\phi_i'(-T_i^*, c_i^*) - \lambda\phi_i(-T_i^*, c_i^*) = -\frac{f(c_i^*)}{1 - F(c_i^*)}\mu(c_i^*). \tag{A.6}$$

This is the first-order condition expressed in equation (9).

Immobile firms. When the firms are immobile by assumption and there is no leakage, we have $\phi_i'(C(c), c) = 0$, $\phi_i(C(c), c) = f(c)$ and $\phi_j(C(c), c) = 0$ for all c. Now the integral $\mu(c_i^*)$ as defined in (A.3) can be simplified to:

$$\mu(c_i^*) = \int_{\underline{c}}^{c_i^*} -\lambda f(c)dc = -\lambda F(c_i^*)$$

Plug this into equation (A.2) to get the optimal externality price with immobile firms:

$$c_i^* = \frac{D}{1+\lambda} + \frac{\mu(c_i^*)}{(1+\lambda)\phi_i(-T_i^*, c_i^*)} = \frac{D}{1+\lambda} - \frac{\lambda}{1+\lambda}\frac{F(c_i^*)}{f(c_i^*)} \equiv c_B \tag{A.7}$$

We denote the externality price with immobile firms, given in equation (A.7), as c_B for future reference.

Mobile firms. We show that mobile firms, characterized by distribution $G(\theta)$, always indicate an optimal externality price that is above c_B we defined above in (A.7). The proof is by contradiction. Assume that $c_i^* \leq c_B$. Use expression (5), $\phi_i(C(c), c) = (1 - G(c))f(c)$ and $\phi_i'(C(c), c) = -g(c)f(c)$, where we denote $G(C(c)) = G(c)$ and $g(C(c)) = g(c)$ for shorthand, to write the first-order condition for c_i^* in (A.1):

$$D - (1+\lambda)c_i^* + \int_{\underline{c}}^{c_i^*} -\Delta_L(c)\frac{g(c)}{1 - G(c_i^*)}\frac{f(c)}{f(c_i^*)} - \lambda\frac{1 - G(c)}{1 - G(c_i^*)}\frac{f(c)}{f(c_i^*)}dc = 0 \tag{A.8}$$

Here, we have denoted the welfare effect of relocation for firms that cut emissions as: $-\Delta_L(c) = \gamma - \lambda T_i^* + D - c - \lambda c_i^*$, where the subindex L refers to "Local" policy. Use the assumption $c_i^* \leq c_B$, implying $D - (1+\lambda)c_i^* \geq \lambda\frac{F(c_i^*)}{f(c_i^*)}$, to write first-order condition (A.8) as the following inequality:

$$\int_{\underline{c}}^{c_i^*} -\Delta_L(c)\frac{g(c)}{1 - G(c_i^*)}\frac{f(c)}{f(c_i^*)} \underbrace{-\lambda\frac{1 - G(c)}{1 - G(c_i^*)}\frac{f(c)}{f(c_i^*)}}_{\equiv A} dc \leq -\lambda\frac{F(c_i^*)}{f(c_i^*)} \tag{A.9}$$

Integrate term A by parts (note, that $C'(c) = 1$), and use $F(\underline{c}) = 0$ to write:

$$A = -\int_{\underline{c}}^{c_i^*} \lambda \frac{1 - G(c)}{1 - G(c_i^*)} \frac{f(c)}{f(c_i^*)} dc = -\lambda \frac{F(c_i^*)}{f(c_i^*)} - \int_{\underline{c}}^{c_i^*} \frac{g(c)}{1 - G(c_i^*)} \frac{F(c)}{f(c_i^*)} dc$$

Using this, inequality (A.9) becomes:

$$\int_{\underline{c}}^{c_i^*} \frac{g(c)}{1 - G(c_i^*)} \frac{f(c)}{f(c_i^*)} \Big(\underbrace{-\Delta_L(c) - \lambda \frac{F(c)}{f(c)}}_{\equiv B} \Big) dc \leq 0 \qquad \text{(A.10)}$$

The inequality holds true if the term B, defined above, is nonpositive. Using the definition of $-\Delta_L$:

$$-\Delta_L(c) = \gamma - \lambda T_i^* + D - c - \lambda c_i^* > D - c - \lambda c_i^* > D - (1 + \lambda) c_i^* \geq \lambda \frac{F(c_i^*)}{f(c_i^*)}$$

The first inequality follows from the fact that $c_j^* \leq c_B$ implies $\mu(c_i^*) \leq -\lambda F(c_i^*) < 0$, which, by first-order condition (A.5), leads to $\gamma - \lambda T > 0$. The second inequality follows from the fact that c_i^* is the upper integral bound in (A.10), and therefore $c \leq c_i^*$. The third inequality follows from $c_i^* \leq c_B$, implying $D - (1 + \lambda) c_i^* \geq \lambda \frac{F(c_i^*)}{f(c_i^*)}$. Term B in (A.10) then writes as:

$$B = -\Delta_L(c) - \lambda \frac{F(c)}{f(c)} > \lambda \frac{F(c_i^*)}{f(c_i^*)} - \lambda \frac{F(c)}{f(c)} \geq 0 \qquad \text{(A.11)}$$

where the last inequality follows from the hazard rate assumption (Assumption 1), and the fact that c_i^* is the upper bound of the integral in (A.10). However, inequality (A.11) leads to a contradiction with (A.10). It must therefore be that, with mobile firms, $c_i^* > c_B$.

This completes the proof of Theorem 1. Q.E.D.

Type-tailored schedule (c observable, θ unobservable)

Derivation of Equation (7). In the text we consider the case where c is observable to the regulator, but θ is not. That is, the regulator chooses $X_i(c)$ and $C_i(c)$ for each c to solve:

$$\max_{X_i(c),C_i(c)} W_i = \int_{\underline{c}}^{\overline{c}} \Big(\gamma + DX_i(c) - cX_i(c) - \lambda(cX_i(c) - C_i(c))\Big)\phi_i(C(c),c)dc$$

where we have used $C_i(c) = cX_i(c) - T_i(c)$. The objective function is linear in $X_i(c)$ so the optimal solution takes a bang-bang form where either $X_i(c) = 0$ or $X_i(c) = 1$. The condition for $X_i(c) = 1$ is:

$$\frac{\partial W_i}{\partial X_i(c)} \geq 0 \Rightarrow (1+\lambda)c \leq D \tag{A.12}$$

and $X_i(c) = 0$ otherwise. By Assumption 1, we show that the objective is monotonic in $C_i(c)$ and that there exists bounded optimal $C_i(c)$. The first-order condition for $C_i(c)$ is:

$$\frac{\partial W_i}{\partial C_i(c)} = 0 \Rightarrow \underbrace{-\Big(\gamma + (D-c)X_i(c) - \lambda T_i(c)\Big)}_{=\Delta(c)}\phi_i'(C(c),c) - \lambda\phi_i(C(c),c) = 0, \tag{A.13}$$

This is the elasticity rule, equation (7) in the main text. Which can be rewritten as

$$\lambda\frac{1-G(C(c))}{-g(C(c))} = \Delta(c).$$

By Assumption 1, the left-hand side is increasing while, by definition, $\Delta(.)$ is strictly decreasing in $C(c)$ for any given c. Thus, the solution to (A.13) exists and is unique.

Properties of the type-tailored schedule (Figure 1). When c is observable:
Property (i) For $c > \frac{D}{1+\lambda}$, the optimal type-tailored compensation schedule is flat:

$$T_i'(c) = 0.$$

To see this, note that when $X_i(c) = 0$, the first-order condition for transfers (A.13) becomes:

$$\lambda\phi(-T_i(c),c) = -\Big(\gamma - \lambda T_i(c)\Big)\phi'(-T_i(c),c)$$

It immediately follows from this and (A.13) that $T_i'(c) = 0$ for $c > D/(1+\lambda)$.

Property (ii) For $c = \frac{D}{1+\lambda}$, the optimal type-tailored compensation jumps discontinuously:

$$\lim_{c\to D/(1+\lambda)^-} T_i(c) - \lim_{c\to D/(1+\lambda)^+} T_i(c) = c_P.$$

To show this, we evaluate the optimal compensation at both sides of $c_P = D/(1+\lambda)$. Denote $T_i^-(c_P) = \lim_{c\to D/(1+\lambda)^-} T_i(c_P)$ and $T_i^+(c_P) = \lim_{c\to D/(1+\lambda)^+} T_i(c_P)$. At c_P the regulator is indifferent between $X_i(c) = 0$ and $X_i(c) = 1$ so condition (A.13) must hold for both of these actions:

$$-\Big(\gamma + (D - c_P) - \lambda T^-(c_P)\Big)\eta_i(c_P - T^-(c_P)) = -\Big(\gamma - \lambda T^+(c_P)\Big)\eta_i(-T^+(c_P))$$

where $\eta_i(\cdot) = \phi_i(\cdot)/\phi_i'(\cdot)$. Plug in $D = (1+\lambda)c_P$ and simplify:

$$\Big(\gamma + \lambda(c_P - T^-(c_P))\Big)\eta_i(c_P - T^-(c_P)) = \Big(\gamma - \lambda T^+(c_P)\Big)\eta_i(-T^+(c_P))$$

It can be verified that this equation is satisfied for $T^-(c_P) - T^+(c_P) = c_P$.

Property (iii) For $c < \frac{D}{1+\lambda}$, the optimal type-tailored compensation satisfies

$$T_i'(c) = \begin{cases} > 0 & \text{iff } \lambda\eta_i'(C_i(c)) > 1 \\ < 0 & \text{iff } \lambda\eta_i'(C_i(c)) < 1 \end{cases}$$

To show this, use $C_i(c) = cX_i(c) - T_i(c)$ and $\eta_i(\cdot) = \phi(\cdot)/\phi'(\cdot)$ to write condition (A.13) as

$$\lambda\eta_i(C_i(c)) = -\Big(\gamma + (D - (1+\lambda)c)X_i(c) + \lambda C_i(c)\Big)$$

Then, differentiate both sides with respect to c:

$$\lambda\eta_i'(C_i(c))C_i'(c) = (1+\lambda)X_i(c) - \lambda C_i'(c)$$
$$\Rightarrow$$
$$C_i'(c) = \frac{(1+\lambda)X_i(c)}{\lambda(\eta_i'(C_i(c)) + 1)}.$$

By $C_i(c) = cX_i(c) - T_i(c)$, condition $T_i'(c) < 0$ is equivalent to $C_i'(c) > X_i(c)$ (recall that $X_i(c) = 1$ for all $c \le \frac{D}{1+\lambda}$):

$$C_i'(c) = \frac{(1+\lambda)X_i(c)}{\lambda(\eta_i'(C_i(c)) + 1)} > X_i(c)$$
$$\Rightarrow$$
$$1 > \lambda\eta_i'(C_i(c))$$

And likewise for $T_i'(c) > 0$.

Proof of Proposition 1.

Proof is by contradiction. Assume that (i) $c_i^* \leq c_P = \frac{D}{1+\lambda}$, implying $\mu(c_i^*) \leq 0$, and (ii) that climate surplus effect dominates the compliance cost effect: $\lambda\big(\eta_i(C(c_P)) - \eta_i(C(c))\big) < D - c - \lambda c_P$. From equation (9) we see that this condition yields:

$$-\big(\gamma - \lambda T_i^*\big)\phi_i'(C(c_i^*), c_i^*) - \lambda\phi_i(C(c_i^*), c_i^*) \geq 0$$

Using definition (6) for the inverse hazard rate, $\eta_i(C(c))$, this can be rewritten as:

$$-\big(\gamma - \lambda T_i^*\big) \leq \lambda\eta_i(C(c_i^*))$$

which can be written by (recall that by definition both $\phi'(C(c), c)$ and $\eta_i(C)$ are negative):

$$-\frac{\gamma - \lambda T_i^*}{\eta_i(C(c))} \geq \lambda\frac{\eta_i(C(c_i^*))}{\eta_i(C(c))} \tag{A.14}$$

Next, we turn attention to the first-order condition (10), which, using definition for $\mu(c_i^*)$ in (8), can be written as:

$$D - (1+\lambda)c_i^* + \int_{\underline{c}}^{c_i^*}\Big(-\frac{\gamma - \lambda T_i^*}{\eta_i(C(c))} - \frac{D - c - \lambda c_i^*}{\eta_i(C(c))} - \lambda\Big)\frac{\phi_i(C(c), c)}{\phi_i(C(c_i^*), c_i^*)}dc = 0 \tag{A.15}$$

And plugging in (A.14), the inequality writes as:

$$D - (1+\lambda)c_i^* + \int_{\underline{c}}^{c_i^*}\Big(\lambda\frac{\eta_i(C(c_i^*))}{\eta_i(C(c))} - \frac{D - c - \lambda c_i^*}{\eta_i(C(c))} - \lambda\Big)\frac{\phi_i(C(c), c)}{\phi_i(C(c_i^*), c_i^*)}dc \leq 0 \tag{A.16}$$

Consider the assumption that climate surplus term dominates:

$$\lambda\big(\eta_i(C(c_P)) - \eta_i(C(c))\big) < D - c - \lambda c_P$$

where we have used the assumption $c_i^* \leq c_P$, the fact that $\eta_i(C)$ is increasing in C and $C'(c) = 1 > 0$.

$$\lambda\big(\eta_i(C(c_i^*)) - \eta_i(C(c))\big) < D - c - \lambda c_i^*$$

$$\Rightarrow$$

$$-\frac{D - c - \lambda c_i^*}{\eta_i(C(c)} > -\lambda\frac{\eta_i(C(c_i^*) - \eta_i(C(c))}{\eta_i(C(c))}$$

Use this in (A.16) to write:

$$D - (1+\lambda)c_i^* + \int_{\underline{c}}^{c_i^*}\underbrace{\Big(\lambda\frac{\eta_i(C(c_i^*))}{\eta_i(C(c))} - \lambda\frac{\eta_i(C(c_i^*) - \eta_i(C(c))}{\eta_i(C(c))} - \lambda\Big)}_{\equiv E}\frac{\phi_i(C(c), c)}{\phi_i(C(c_i^*), c_i^*)}dc < 0$$

By simplifying, it can be seen that term E equals zero, and the inequality becomes:

$$D - (1+\lambda)c_i^* < 0 \quad\Rightarrow\quad c_i^* > \frac{D}{1+\lambda} \tag{A.17}$$

However, (A.17) is a contradiction to the starting assumption that $c_i^* \leq \frac{D}{1+\lambda}$. Thus, the assumption $\lambda\big(\eta_i(C(c_i^*)) - \eta_i(C(c))\big) < D - c - \lambda c_i^*$ must imply $c_i^* > c_P = \frac{D}{1+\lambda}$. Q.E.D.

Proof of Proposition 2.

The regulator is constrained to use a purely global mechanism, that is, the same externality price in both countries: $c_i^* = c_j^* = c^*$. Under the global cap, the firms' net cost is the same regardless of its type: $C(c) = C_i(c) - C_j(c) = c - c^* - T_i^* - (c - c^* - T_j^*) = T_j^* - T_i^*$ regardless of whether the firm cuts emissions or not. The regulator is left to find the compensations, T_i^* and T_j^*, and the global price level, c^*, to maximize the social welfare:

$$W_i = \int_{\underline{c}}^{c^*} \Big(\gamma + D - c - \lambda(T_i^* + c^*)\Big)\phi_i(T_j^* - T_i^*, c) + \Big(D - (1+\lambda)\big(T_j^* + c^*\big)\Big)\phi_j(T_j^* - T_i^*, c)dc +$$
$$\int_{c^*}^{\overline{c}} \Big(\gamma - \lambda T_i^*\Big)\phi_i(T_j^* - T_i^*, c) - (1+\lambda)T_j^*\phi_j(T_j^* - T_i^*, c)dc$$

such that $C_j(c) = -T_j^* \leq 0$ for all c. Noting that W_i is decreasing in T_j^*, it follows that the regulator would optimally set the foreign base compensation level (paid to firms who do not cut) to zero: $T_j^* = 0$. Take the first-order condition with respect to the global price, c^*:

$$\frac{\partial W_i}{\partial c^*} = \Big(D - (1+\lambda)c^*\Big)\phi_i(C(c^*), c^*) + \Big(D - (1+\lambda)c^*\Big)\phi_j(C(c^*), c^*) +$$
$$\int_{\underline{c}}^{c^*} -\lambda\phi_i(C(c^*), c^*) - (1+\lambda)\phi_j(C(c^*), c^*)dc = 0 \tag{A.18}$$

Next, observe that (5) implies $\phi_i(C(c), c) + \phi_j(C(c), c) = f(c)$, and that $\phi_j(C(c), c) = G(C(c))f(c)$. Use this to write equation (A.18) as:

$$\big(D - (1+\lambda)c^*\big)f(c^*) - \int_{\underline{c}}^{c^*} \big(G(c^*) + \lambda\big)f(c)dc = 0$$
$$\Rightarrow$$
$$c^* = \frac{D}{1+\lambda} - \frac{G(c^*) + \lambda}{1+\lambda}\frac{F(c^*)}{f(c^*)}$$

This is the uniform emission price in the purely-global mechanism. Q.E.D.

Proof of Theorem 2 (Global mechanism).

Lemmas 1-2 continue to hold. We begin by deriving a global counterpart for Lemma 3:

Lemma 4. *In the optimal global mechanism ($C_j(c) \leq 0$), actions take a bang-bang form where $X_i(c) = \{0, 1\}$ and $X_j(c) = \{0, 1\}$*

Proof. Objective function (1) can be written as:

$$\max_{X_i(c),C_i(c),X_j(c),C_j(c)} \int_{\underline{c}}^{\overline{c}} \Big(\gamma + DX_i(c) - (1+\lambda)cX_i(c) + \lambda C_i(c)\Big)\phi_i\big(C_i(c) - C_j(c), c\big) + \Big(DX_j(c) - (1+\lambda)cX_j(c) + (1+\lambda)C_j(c)\Big)\phi_j\big(C_i(c) - C_j(c), c\big)dc$$

s.t. $C_k'(c) = -X_k(c)$ holds for all c and $k = i, j$. Denoting the co-state variables by $\mu_i(c)$ and $\mu_j(c)$, the Hamiltonian for this problem reads:

$$\mathcal{H} = \big(\gamma + DX_i(c) - (1+\lambda)cX_i(c) + \lambda C_i(c)\big)\phi_i(C_i(c) - C_j(c), c) + \big(DX_j(c) - (1+\lambda)cX_j(c) + (1+\lambda)C_j(c)\big)\phi_j\big(C_i(c) - C_j(c), c\big) - \mu_i(c)X_i(c) - \mu_j(c)X_j(c)$$

The Hamiltonian is linear in the controls $X_i(c)$ and $X_j(c)$, and the necessary conditions for optimality imply that $X_i(c)$ and $X_j(c)$ take bang-bang forms: $X_i(c) = \{0, 1\}$ and $X_j(c) = \{0, 1\}$. Q.E.D.

Based on Lemmas 1,2 and 4, the mechanism boils down to finding a tuple $(c_i^*, c_j^*, T_i^*, T_j^*)$ that defines $M_i(c)$ and $M_j(c)$ through base compensations and thresholds for cuts at home and abroad. We prove Theorem 2 in two steps. Denote the benchmark for pure global mechanism as defined in Proposition 2: $c^* = \frac{D}{1+\lambda} - \frac{G(c^*)+\lambda}{1+\lambda}\frac{F(c^*)}{f(c^*)}$. To begin with, we consider the case that $c_j^* < c_i^*$ and show that this implies $c_i^* > c^*$ (Step 1). Second, we consider the case $c_j^* \geq c_i^*$ and show that this leads to a contradiction (Step 2).

Step 1. We begin by considering the case $c_j^* < c_i^*$. In that case, the maximization problem can be written as:

$$\max_{c_i^*,c_j^*,T_i^*,T_j^*} W_i = \int_{\underline{c}}^{c_j^*} \big(\gamma + D - c - \lambda(T_i^* + c_i^*)\big)\phi_i(C(c), c) - \big(D - (1+\lambda)(T_j^* + c_j^*)\big)\phi_j(C(c), c)dc + \int_{c_j^*}^{c_i^*} \big(\gamma + D - c - \lambda(T_i^* + c_i^*)\big)\phi_i(C(c), c) - (1+\lambda)T_j^*\phi_j(C(c), c)dc + \int_{c_i^*}^{\overline{c}} \big(\gamma - \lambda T_i^*\big)\phi_i(C(c), c) - (1+\lambda)T_j^*\phi_j(C(c), c)dc$$

such that $C_j(c) \leq 0$ for all c. The first integral denotes firms under the global cap c_j^*, the second integral are the firms under only a local price c_i^* and the last integral the firms who do not cut in either location. It can be seen that W_i is always decreasing in T_j^*, which, together with the constraint $C_j(c) \leq 0$, guarantees that $T_j^* = 0$: Participation constraint prevents regulator from taxing the firms abroad, but it gains nothing from subsidizing them either.

To simplify notation, we use the definition of the net welfare effect of relocation, as defined in equation (2). For firms under the global cap ($c \leq c_j^*$), we have $-\Delta_G(c) = \gamma - \lambda T_i^* - c - \lambda c_i^* + (1+\lambda)c_j^*$, and, for those under the local cap ($c_j^* < c \leq c_i^*$), we have $-\Delta_L(c) = \gamma - \lambda T_i^* + D - c - \lambda c_i^*$. When firms cut in both regimes ($c \leq c_j^*$), their net cost, $C(c) = C_i(c) - C_j(c) = c_j^* - c_i^* - T_i^*$, is independent of their type c. When firms cut only when they stay ($c_j^* < c \leq c_i^*$), their net cost is: $C(c) = c - c_i^* - T_i^*$, and finally, when firms cut in neither location ($c_i^* < c$) their net cost, $C(c) = -T_i^*$, is again independent of c.

We aim to show $c_j^* < c_i^* \Rightarrow c_i^* > c^*$ so that the local externality price is above the uniform-price benchmark. The proof is by contradiction: assume that $c_i^* \leq c^*$ which, by $c_j^* < c_i^*$, also implies that $c_j^* < c^*$. To show contradiction, we combine the first-order conditions for c_j^* (Step 1a), T_i^* (Step 1b) and c_i^* (Step 1c). The contradiction is shown at the end of Step 1c.

Step 1a. The first-order condition with respect to the global cap, c_j^*, is:

$$(D - (1+\lambda)c_j^*)\phi_j(C(c_j^*), c_j^*) + \int_{\underline{c}}^{c_j^*} \Big[\Delta_G(c)\phi_j'(C(c_j^*), c) - (1+\lambda)\phi_j(C(c_j^*), c)\Big] dc = 0 \tag{A.19}$$

We use expressions (5) to write $\phi_j(C(c), c) = G(C(c))f(c)$, and $\phi_j'(C(c), c) = g(c)f(c)$ where, to save notation, we write $G(C(c)) = G(c)$ and $g(C(c)) = g(c)$. The first-order condition (A.19) becomes:

$$D - (1+\lambda)c_j^* + \int_{\underline{c}}^{c_j^*} \Big(\Delta_G(c)\frac{g(c_j^*)}{G(c_j^*)}\frac{f(c)}{f(c_j^*)} - (1+\lambda)\frac{f(c)}{f(c_j^*)}\Big) dc = 0$$

Now, we use the condition that $c_j^* < c_i^* < c^*$:

$$\begin{aligned}\int_{\underline{c}}^{c_j^*} \Delta_G(c)\frac{g(c_j^*)}{G(c_j^*)}\frac{f(c)}{f(c_j^*)} dc - (1+\lambda)\frac{F(c_j^*)}{f(c_j^*)} = -\big(D - (1+\lambda)c_j^*\big) < -\big(D - (1+\lambda)c^*\big) = \\ -\big(G(c^*) + \lambda\big)\frac{F(c^*)}{f(c^*)} < -\big(G(c_i^*) + \lambda\big)\frac{F(c_j^*)}{f(c_j^*)}\end{aligned} \tag{A.20}$$

Where the first inequality follows from the assumption $c_j^* < c^*$, the second inequality follows from the definition of c^* and the last inequality follows from the fact that both $G(c)$ and $F(c)/f(c)$ are increasing in c. The inequality (A.20) above can be rewritten as:

$$\int_{\underline{c}}^{c_j^*} -\Delta_G(c)\frac{g(c_j^*)}{1 - G(c_i^*)}\frac{f(c)}{f(c_i^*)} dc > -G(c_j^*)\frac{F(c_j^*)}{f(c_i^*)} > -G(c_i^*)\frac{F(c_i^*)}{f(c_i^*)} \tag{A.21}$$

where for the last step we used the fact that $G(c)$ and $F(c)/f(c)$ are increasing and $c_j^* < c_i^*$.

Step 1b. The first-order condition for the optimal T_i^*, using with the first-order condition with respect to c_i^* (as shown in eq. (A.24)), can be written as:

$$\Big(-[\gamma-\lambda T_i^*]\phi_i'(C(c_i^*),c_i^*)-\lambda\phi_i(C(c_i^*),c_i^*)\Big)\frac{1-F(c_i^*)}{f(c_i^*)}=\big(D-(1+\lambda)c_i^*\big)\phi_i(C(c_i^*),c_i^*)$$

We, again, use expressions $\phi_i(C(c),c))=(1-G(c))f(c)$, and $\phi_i'(C(c),c)=-g(c)f(c)$ to write:

$$\Big([\gamma-\lambda T_i^*]\frac{g(c_i^*)}{1-G(c_i^*)}-\lambda\Big)\frac{1-F(c_i^*)}{f(c_i^*)}=D-(1+\lambda)c_i^*>0$$

where the inequality follows from $c_i^*<c^*<\frac{D}{1+\lambda}$. It immediately follows that:

$$\gamma-\lambda T_i^*>0 \tag{A.22}$$

Write the definition of $\Delta_L(c)$,

$$-\Delta_L(c)=\gamma-\lambda T_i^*+D-c-\lambda c_i^*>D-(1+\lambda)c_i^*\geq(G(c^*)+\lambda)\frac{F(c^*)}{f(c^*)}>\lambda\frac{F(c_i^*)}{f(c_i^*)}, \tag{A.23}$$

where: for the first inequality we have used (A.22) and the fact that $c<c_i^*$; for the second inequality we have used $c_i^*<c^*$; and for the last inequality we have used $G(c^*)>0$, the fact that $F(c)/f(c)$ is increasing in c and $c_i^*<c^*$.

Step 1c. Next, we turn focus on the first-order condition with respect to the local price c_i^*:

$$\begin{aligned}D-(1+\lambda)c_i^*+\int_{\underline{c}}^{c_j^*}\Delta_G(c)\frac{\phi_i'(C(c_j^*),c)}{\phi_i(C(c_i^*),c_i^*)}-\lambda\frac{\phi_i(C(c_j^*),c)}{\phi_i(C(c_i^*),c_i^*)}dc+\\ \int_{c_j^*}^{c_i^*}\Delta_L(c)\frac{\phi_i'(C(c),c)}{\phi_i(C(c_i^*),c_i^*)}-\lambda\frac{\phi_i(C(c),c)}{\phi_i(C(c_i^*),c_i^*)}dc=0\end{aligned}$$

Use expressions $\phi_i(C(c),c))=(1-G(c))f(c)$, and $\phi_i'(C(c),c)=-g(c)f(c)$ to write:

$$\begin{aligned}D-(1+\lambda)c_i^*+\int_{\underline{c}}^{c_j^*}-\Delta_G(c)\frac{g(c_j^*)}{1-G(c_i^*)}\frac{f(c)}{f(c_i^*)}-\lambda\frac{1-G(c_j^*)}{1-G(c_i^*)}\frac{f(c)}{f(c_i^*)}dc+\\ \int_{c_j^*}^{c_i^*}-\Delta_L(c)\frac{g(c)}{1-G(c_i^*)}\frac{f(c)}{f(c_i^*)}-\lambda\frac{1-G(c)}{1-G(c_i^*)}\frac{f(c)}{f(c_i^*)}dc=0\end{aligned} \tag{A.24}$$

Using the assumption $c_i^*\leq c^*$, implying: $D-(1+\lambda)c_i^*\geq(G(c^*)+\lambda)\frac{F(c^*)}{f(c^*)}$, and simplifying, we get:

$$\begin{aligned}\int_{\underline{c}}^{c_j^*}-\Delta_G(c)\frac{g(c_j^*)}{1-G(c_i^*)}\frac{f(c)}{f(c_i^*)}dc-\lambda\frac{1-G(c_j^*)}{1-G(c_i^*)}\frac{F(c_j^*)}{f(c_i^*)}+\\ \int_{c_j^*}^{c_i^*}-\Delta_L(c)\frac{g(c)}{1-G(c_i^*)}\frac{f(c)}{f(c_i^*)}\underbrace{-\lambda\frac{1-G(c)}{1-G(c_i^*)}\frac{f(c)}{f(c_i^*)}dc}_{\equiv H}\leq-\big(G(c^*)+\lambda\big)\frac{F(c^*)}{f(c^*)}\end{aligned} \tag{A.25}$$

Integrate term H in equation (A.25) by parts (Note, that $dG(C(c)) = C'(c)g(C(c)) = g(C(c))$, because $C'(c) = 1$):

$$H = -\int_{c_j^*}^{c_i^*} \lambda \frac{1-G(c)}{1-G(c_i^*)} \frac{f(c)}{f(c_i^*)} dc = -\lambda \frac{F(c_i^*)}{f(c_i^*)} + \lambda \frac{1-G(c_j^*)}{1-G(c_i^*)} \frac{F(c_j^*)}{f(c_i^*)} - \int_{\underline{c}}^{c_i^*} \lambda \frac{g(c)}{1-G(c_i^*)} \frac{F(c)}{f(c_i^*)} dc$$

Use this, together with (A.21), to write the left-hand side of equation (A.25) as:

$$-\Big(G(c_i^*) + \lambda\Big) \frac{F(c_i^*)}{f(c_i^*)} + \int_{c_j^*}^{c_i^*} \frac{g(c)}{f(c_i^*)} \frac{1}{1-G(c_i^*)} \Big(\underbrace{-\Delta_L f(c) - \lambda F(c)}_{\equiv I}\Big) dc > -\big(G(c^*) + \lambda\big) \frac{F(c^*)}{f(c^*)} \quad \text{(A.26)}$$

The inequality follows from, firstly, the fact that $G(c)$ and $F(c)/f(c)$ are increasing in c and secondly, because the second term is positive since, as can be seen by using the definition for $\Delta_L(c)$ in (A.23):

$$I = -\Delta_L(c) f(c) - \lambda F(c) > \lambda \frac{F(c_i^*)}{f(c_i^*)} f(c) - \lambda F(c) \geq 0, \quad \text{(A.27)}$$

However, there is contradiction between (A.25) and (A.26). This proves that if $c_i^* > c_j^*$, we must have $c_i^* > c^*$.

Step 2. Let us next assume $c_i^* \leq c_j^*$. This is shown to lead to a contradiction both when $\gamma - \lambda T_i^* > 0$ (Step 2a) and $\gamma - \lambda T_i^* \leq 0$ (Step 2b). With $c_i^* \leq c_j^*$ the maximization problem is written as:

$$\begin{aligned}\max_{c_i^*, c_j^*, T_i^*, T_j^*} W_i = &\int_{\underline{c}}^{c_i^*} \big(\gamma + D - c - \lambda(T_i^* + c_i^*)\big)\phi_i(C(c), c) + \big(D - (1+\lambda)(T_j^* + c_j^*)\big)\phi_j(C(c), c) dc + \\ &\int_{c_i^*}^{c_j^*} \big(\gamma - \lambda T_i^*\big)\phi_i(C(c), c) + \big(D - (1+\lambda)(T_j^* + c_j^*)\big)\phi_j(C(c), c) dc + \\ &\int_{c_j^*}^{\overline{c}} \big(\gamma - \lambda T_i^*\big)\phi_i(C(c), c) - (1+\lambda) T_j^* \phi_j(C(c), c) dc\end{aligned}$$

such that $C_j(c) \leq 0$. As before, we can immediately conclude that $T_j^* = 0$ optimally. In this case, firms cut only in j if $c_i^* \leq c \leq c_j^*$. The welfare effects of the relocation under the foreign cap defined as $-\Delta_F(c) = \gamma - \lambda T_i^* - D + (1+\lambda)c_j^*$, for $c_i^* \leq c \leq c_j^*$. In contrast to Step 1, the net cost of cutting under the foreign price ($c_i^* < c \leq c_j^*$): $C(c) = -c + c_j^* - T_i^*$, implying $C'(c) = -1$.

Step 2a. Assume first that $\gamma - \lambda T_i^* > 0$. In that case, we can write:

$$-\Delta_G(c) = \gamma - \lambda T_i^* + (1+\lambda)c_j^* - c - \lambda c_i^* \geq -\Delta_G(c_i^*) = \gamma - \lambda T_i^* + (1+\lambda)c_j^* - (1+\lambda)c_i^* > 0 \quad \text{(A.28)}$$

where the first inequality follows from the fact that c_i^* is the upper integral bound ($c \leq c_i^*$) and the last inequality follows from $\gamma - \lambda T_i^* > 0$ and $c_j^* \geq c_i^*$. The first-order condition for c_i^* can be written as:

$$-D + (1+\lambda)c_i^* = \int_{\underline{c}}^{c_i^*} -\Delta_G(c)\frac{g(c_i^*)}{1-G(c_i^*)}\frac{f(c)}{f(c_i^*)}dc - \lambda\frac{F(c_i^*)}{f(c_i^*)} > -\lambda\frac{F(c_i^*)}{f(c_i^*)} \tag{A.29}$$

where the last inequality follows from $-\Delta_G(c) > 0$, as stated in (A.28). Using this, we can write the definition for $\Delta_F(c)$ as:

$$-\Delta_F(c) = \gamma - \lambda T_i^* - D + (1+\lambda)c_j^* \geq -D + (1+\lambda)c_i^* > -\lambda\frac{F(c_i^*)}{f(c_i^*)} \tag{A.30}$$

where the first inequality follows from using (A.28), and the for the second inequality we have used (A.29). The first-order condition for the foreign price, c_j^*, can be written as:

$$\begin{aligned} D - (1+\lambda)c_j^* + \int_{\underline{c}}^{c_i^*} \Delta_G(c)\frac{g(c_i^*)}{G(c_j^*)}\frac{f(c)}{f(c_j^*)}dc - (1+\lambda)\frac{G(c_i^*)}{G(c_j^*)}\frac{F(c_i^*)}{f(c_j^*)} + \\ \int_{c_i^*}^{c_j^*} \Delta_F(c)\frac{g(c)}{G(c_j^*)}\frac{f(c)}{f(c_j^*)} - (1+\lambda)\frac{G(c)}{G(c_j^*)}\frac{f(c)}{f(c_j^*)}dc = 0. \end{aligned} \tag{A.31}$$

From the assumption $c_j^* \geq c_i^*$ it follows that we must have:

$$-\big(D - (1+\lambda)c_j^*\big) \geq -\big(D - (1+\lambda)c_i^*\big) \tag{A.32}$$

Plug the first-order condition for c_i^* in (A.29) and c_j^* in equation (A.31) into (A.32), to write:

$$\begin{aligned} &\int_{\underline{c}}^{c_i^*} \Delta_G(c)\frac{g(c_i^*)}{G(c_j^*)}\frac{f(c)}{f(c_j^*)}dc - (1+\lambda)\frac{G(c_i^*)}{G(c_j^*)}\frac{F(c_i^*)}{f(c_j^*)} + \int_{c_i^*}^{c_j^*} \Delta_F(c)\frac{g(c)}{G(c_j^*)}\frac{f(c)}{f(c_j^*)} \underbrace{-(1+\lambda)\frac{G(c)}{G(c_j^*)}\frac{f(c)}{f(c_j^*)}}_{\equiv V} dc \geq \\ &\int_{\underline{c}}^{c_i^*} -\Delta_G(c)\frac{g(c_i^*)}{1-G(c_i^*)}\frac{f(c)}{f(c_i^*)}dc - \lambda\frac{F(c_i^*)}{f(c_i^*)} \end{aligned} \tag{A.33}$$

Integrate term V by parts, noting that $C'(c) = -1$, to write:

$$V = -(1+\lambda)\frac{G(c)}{G(c_j^*)}\frac{f(c)}{f(c_j^*)} = -(1+\lambda)\frac{F(c_j^*)}{f(c_j^*)} + (1+\lambda)\frac{G(c_i^*)}{G(c_j^*)}\frac{F(c_i^*)}{f(c_j^*)} - \int_{c_i^*}^{c_j^*}(1+\lambda)\frac{g(c)}{G(c_j^*)}\frac{F(c)}{f(c_j^*)}dc$$

Plugging term V into equation (A.33) can be written as:

$$\begin{aligned} &\int_{\underline{c}}^{c_i^*} \Delta_G(c)g(c_j^*)f(c)\Big(\frac{1}{G(c_j^*)}\frac{1}{f(c_j^*)} + \frac{1}{1-G(c_i^*)}\frac{1}{f(c_i^*)}\Big)dc \\ &+ \int_{c_i^*}^{c_j^*} \frac{g(c)}{G(c_j^*)}\frac{1}{f(c_j^*)}\Big(\underbrace{\Delta_F(c)f(c) - (1+\lambda)F(c)}_{\equiv J}\Big)dc - \Big((1+\lambda)\frac{F(c_j^*)}{f(c_j^*)} - \lambda\frac{F(c_i^*)}{f(c_i^*)}\Big) \geq 0 \end{aligned} \tag{A.34}$$

From (A.28) we know that the first term is non-positive. We can also see that term J is strictly negative as:

$$J = \Delta_F(c)f(c) - (1+\lambda)F(c) < \lambda\frac{F(c_i^*)}{f(c_i^*)}f(c) - (1+\lambda)F(c) \leq 0$$

where the first inequality follows from (A.30) and the second inequality follows from the fact that c_i^* is the lower integral bound and $F(c)/f(c)$ is increasing in its argument by Assumption 1. Last, we can see that this and the fact that $c_i^* > c_j^*$ imply that the last term in equation (A.34) is strictly negative. The left-hand side of (A.34) is therefore strictly negative, however, this is a contradiction to the inequality in (A.34). Therefore, $\gamma - \lambda T_i^* > 0$ implies that $c_i^* > c_j^*$.

Step 2b. To complete the proof, we are left to show that the case $\gamma - \lambda T_i^* \leq 0$ also leads to a contradiction. We move to first-order condition for the baseline compensation, T_i^*:

$$\Big([\gamma - \lambda T_i^*]\frac{g(c_j^*)}{1-G(c_j^*)} - \lambda\Big)\frac{1-F(c_j^*)}{f(c_j^*)} =$$

$$\int_{\underline{c}}^{c_i^*}\Delta_G(c)\frac{g(c_i^*)}{1-G(c_j^*)}\frac{f(c)}{f(c_j^*)}dc + \lambda\frac{1-G(c_i^*)}{1-G(c_j^*)}\frac{F(c_i^*)}{f(c_j^*)} + \int_{c_i^*}^{c_j^*}\Delta_F(c)\frac{g(c)}{1-G(c_j^*)}\frac{f(c)}{f(c_j^*)} + \lambda\frac{1-G(c)}{1-G(c_j^*)}\frac{f(c)}{f(c_j^*)}dc$$

We can notice that $\gamma - \lambda T_i^* \leq 0$ implies that the right-hand side is negative:

$$\int_{\underline{c}}^{c_i^*}\Delta_G(c)\frac{g(c_i^*)}{1-G(c_j^*)}\frac{f(c)}{f(c_j^*)}dc + \lambda\frac{1-G(c_i^*)}{1-G(c_j^*)}\frac{F(c_i^*)}{f(c_j^*)} + \int_{c_i^*}^{c_j^*}\Delta_F(c)\frac{g(c)}{1-G(c_j^*)}\frac{f(c)}{f(c_j^*)} + \lambda\frac{1-G(c)}{1-G(c_j^*)}\frac{f(c)}{f(c_j^*)}dc < 0. \tag{A.35}$$

First, note that the sum of the second and the fourth term in (A.35) is always strictly positive, and a necessary condition for the inequality to hold is therefore:

$$\int_{\underline{c}}^{c_i^*}\Delta_G(c)g(c_i^*)f(c)dc + \int_{c_i^*}^{c_j^*}\Delta_F(c)g(c)f(c)dc < 0 \tag{A.36}$$

By reorganizing the terms, the inequality (A.35) can be written as:

$$\int_{\underline{c}}^{c_i^*}\Delta_G(c)g(c_i^*)f(c)dc + \lambda(1-G(c_i^*))F(c_i^*) + \int_{c_i^*}^{c_j^*}\Delta_F(c)g(c)f(c) + \lambda(1-G(c))f(c)dc < 0$$

$$\Rightarrow$$

$$\int_{\underline{c}}^{c_i^*}\Delta_G(c)g(c_i^*)f(c)dc - \lambda G(c_i^*)F(c_i^*) + \int_{c_i^*}^{c_j^*}\Delta_F(c)g(c)f(c) - \lambda G(c)f(c)dc < -\lambda F(c_i^*) - \int_{c_i^*}^{c_j^*}\lambda f(c)dc$$

$$\Rightarrow$$

$$\int_{\underline{c}}^{c_i^*}\Delta_G(c)g(c_i^*)f(c)dc - \lambda G(c_i^*)F(c_i^*) + \int_{c_i^*}^{c_j^*}\Delta_F(c)g(c)f(c) - \lambda G(c)f(c)dc < -\lambda F(c_j^*) \tag{A.37}$$

Write the first-order condition for c_j^*, given in (A.31), as:

$$-\big(D - (1+\lambda)c_j^*\big) =$$

$$\int_{\underline{c}}^{c_i^*} \Delta_G(c)\frac{g(c_i^*)}{G(c_j^*)}\frac{f(c)}{f(c_j^*)}dc - (1+\lambda)\frac{G(c_i^*)}{G(c_j^*)}\frac{F(c_i^*)}{f(c_j^*)} + \int_{c_i^*}^{c_j^*} \Delta_F(c)\frac{g(c)}{G(c_j^*)}\frac{f(c)}{f(c_j^*)} - (1+\lambda)\frac{G(c)}{G(c_j^*)}\frac{f(c)}{f(c_j^*)}dc <$$

$$-\frac{G(c_i^*)}{G(c_j^*)}\frac{F(c_i^*)}{f(c_j^*)} \underbrace{-\int_{c_i^*}^{c_j^*}\frac{G(c)}{G(c_j^*)}\frac{f(c)}{f(c_j^*)}dc}_{\equiv K} - \lambda\frac{1}{G(c_j^*)}\frac{F(c_j^*)}{f(c_j^*)} \tag{A.38}$$

where for the inequality we have used (A.37). Integrate the term in the K by parts (note: $C'(c) = -1$ when $c_i^* \leq c \leq c_j^*$).

$$K = -\int_{c_i^*}^{c_j^*}\frac{G(c)}{G(c_j^*)}\frac{f(c)}{f(c_j^*)}dc = -\frac{F(c_j^*)}{f(c_j^*)} + \frac{G(c_i^*)}{G(c_j^*)}\frac{F(c_i^*)}{f(c_j^*)} - \int_{c_i^*}^{c_j^*}\frac{g(c)}{G(c_j^*)}\frac{F(c)}{f(c_j^*)}dc$$

Plug this in to write the inequality in (A.38) as:

$$-\left(D-(1+\lambda)c_j^*\right) \leq -\frac{F(c_j^*)}{f(c_j^*)} - \lambda\frac{1}{G(c_j^*)}\frac{F(c_j^*)}{f(c_j^*)} - \int_{c_i^*}^{c_j^*}\frac{g(c)}{G(c_j^*)}\frac{F(c)}{f(c_j^*)}dc \leq -(1+\lambda)\frac{F(c_j^*)}{f(c_j^*)} \tag{A.39}$$

where the last inequality follows from $G(c_j^*) \in (0,1]$. From the assumption $c_i^* \leq c_j^*$, it follows, using equations (A.32) and (A.39), that:

$$-\left(D-(1+\lambda)c_i^*\right) \leq -\left(D-(1+\lambda)c_j^*\right) \leq -(1+\lambda)\frac{F(c_j^*)}{f(c_j^*)} < 0 \tag{A.40}$$

$$\Rightarrow$$

$$\int_{\underline{c}}^{c_i^*} -\Delta_G(c)\frac{g(c_i^*)}{1-G(c_i^*)}\frac{f(c)}{f(c_i^*)}dc \leq \lambda\frac{F(c_i^*)}{f(c_i^*)} - (1+\lambda)\frac{F(c_j^*)}{f(c_j^*)} < 0 \tag{A.41}$$

where, for the last inequality, we have used the fact that $F(c)/f(c)$ is increasing and $c_j^* > c_i^*$. Use the definition of $\Delta_F(c)$:

$$-\Delta_F(c) = \gamma - \lambda T_i^* - D + (1+\lambda)c_j^* \leq 0 \tag{A.42}$$

Where the first inequality follows from the assumption $\gamma - \lambda T_i^* \leq 0$ and using (A.39). Equations (A.41) and (A.42) together imply a contradiction to equation (A.36) when $\gamma - \lambda T_i^* \leq 0$: equation (A.41) shows that the first term in (A.36) must be positive, and equation (A.42) shows that the second term in (A.36) must be positive, implying a contradiction. This completes the proof: Steps 2a and 2b together prove that $c_i^* > c_j^*$. Q.E.D.

Proof of Proposition 3.

Lemmas 1-3 continue to hold if there is perfect correlation between c and θ: The optimal policy is a threshold policy, with $T_i(c) = T_i^*$ for $c > c_i^*$ and $T_i(c) = T_i^* + c_i^*$ for $c \leq c_i^*$. In the general model with full support for θ, some firms move for any given c. In contrast, with perfect correlation, moving firms either have $c > c_i^*$ or $c \leq c_i^*$. Logically, for leaving firms with $c > c_i^*$, we have $-T_i^* > b + kc$, and so

$$c < -\frac{b + T_i^*}{k} \text{ for } k > 0 \tag{A.43a}$$

$$c > -\frac{b + T_i^*}{k} \text{ for } k < 0. \tag{A.43b}$$

Respectively, for leaving firms $c < c_i^*$, we have $c - T_i^* - c_i^* > b + kc$, which leads to

$$c > \frac{b + T_i^* + c_i^*}{1 - k} \text{ for } k < 1 \tag{A.44a}$$

$$c < \frac{b + T_i^* + c_i^*}{1 - k} \text{ for } k > 1. \tag{A.44b}$$

For interior outcomes, we can find three separate cases: (1) when $k < 0$ only condition (A.43b) matters; (2) when $k > 1$ only condition (A.44b) matters; and (3) when $0 < k < 1$ both conditions (A.43a) and (A.44a) are relevant. To prove the Proposition, we analyze these three possibilities separately.

Case 1: $k < 0$**.** Costs c can be divided into three intervals: Firms cut emissions if $c \leq c_i^*$, stay but do not cut emissions when $c_i^* < c \leq c' = -\frac{b+T_i^*}{k}$, and leave when $c > c' = -\frac{b+T_i^*}{k}$, where threshold c' is given by condition (A.43b). The social welfare function can be written as:

$$\max_{c_i^*, T_i^*} \int_{\underline{c}}^{c_i^*} \Big(\gamma + (D - c) - \lambda(T_i^* + c_i^*)\Big) f(c)dc + \int_{c_i^*}^{-\frac{b+T_i^*}{k}} \Big(\gamma - \lambda T_i^*\Big) f(c)dc$$

Take the first-order condition with respect to c_i^*:

$$\Big(D - (1 + \lambda)c_i^*\Big) f(c_i^*) - \int_{\underline{c}}^{c_i^*} \lambda f(c)dc = 0$$

$$\Rightarrow$$

$$c_i^* = \frac{D}{1 + \lambda} - \frac{\lambda}{1 + \lambda} \frac{F(c_i^*)}{f(c_i^*)}, \tag{A.45}$$

which is the expression in Proposition 3.

Case 2: $k > 1$**.** For strong positive correlation, condition (A.44b) is holding and firms leave if $c \leq c' = \frac{b+T_i^*+c_i^*}{1-k}$, stay and cut emissions if $c' < c \leq c_i^*$, and stay without cutting emissions if $c > c_i^*$. The social welfare function can be written as:

$$\max_{c_i^*, T_i^*} \int_{\frac{b+T_i^*+c_i^*}{1-k}}^{c_i^*} \Big(\gamma + (D-c) - \lambda(T_i^* + c_i^*)\Big) f(c)dc + \int_{c_i^*}^{\bar{c}} \Big(\gamma - \lambda T_i^*\Big) f(c)dc$$

Take the first-order condition with respect to c_i^*:

$$\Big(D - (1+\lambda)c_i^*\Big) f(c_i^*) - \frac{1}{1-k} \underbrace{\Big(\gamma + (D-c') - \lambda(T_i^* + c_i^*)\Big)}_{-\Delta(c')} f(c') - \int_{c'}^{c_i^*} \lambda f(c)dc = 0$$

$$\Rightarrow \Big(D - (1+\lambda)c_i^*\Big) f(c_i^*) = -\frac{1}{1-k}\Delta(c')f(c') + \lambda\Big(F(c_i^*) - F(c')\Big). \tag{A.46}$$

Then, the first-order condition with respect to T_i^*:

$$-\frac{1}{1-k} \underbrace{\Big(\gamma + (D-c') - \lambda(T_i^* + c_i^*)\Big)}_{-\Delta(c')} f(c') - \lambda \int_{c'}^{\bar{c}} f(c)dc = 0$$

$$\Rightarrow \Delta(c')f(c') = (1-k)\lambda\Big(1 - F(c')\Big). \tag{A.47}$$

Compensation is chosen so that it balances the marginal damages when a firm of type c' moves (left-hand side) and the overcompensation to all the remaining firms (right-hand side). Plugging equation (A.47) into equation (A.46), we can solve for c_i^* as follows:

$$c_i^* = \frac{D}{1+\lambda} + \frac{\lambda}{1+\lambda} \frac{1 - F(c_i^*)}{f(c_i^*)} \tag{A.48}$$

Since the last term is positive, the emissions price is distorted upwards above the first-best level $D/(1+\lambda)$ and, trivially, also above the benchmark where firms are immobile by assumption.

Case 3: $0 < k < 1$**.** For weak positive correlation, both condition (A.43a) and (A.44a) are relevant. Firms then leave when $\frac{b+T_i^*+c_i^*}{1-k} = c'' < c < c' = -\frac{b+T_i^*}{k}$. In this case, the firms with abatement costs "in the middle" leave, creating two extensive margins: Firms stay and cut emissions if $c \leq c''$ and stay but do not cut when $c \geq c'$. The social welfare function can be written as:

$$\max_{c_i^*, T_i^*} \int_{\underline{c}}^{\frac{b+T_i^*+c_i^*}{1-k}} \Big(\gamma + (D-c) - \lambda(T_i^* + c_i^*)\Big) f(c)dc + \int_{-\frac{b+T_i^*}{k}}^{\bar{c}} \Big(\gamma - \lambda T_i^*\Big) f(c)dc$$

where the first integral denotes firms that cut emissions, and the second part firms that stay but do not cut. Take the first-order condition with respect to c_i^*:

$$\frac{1}{1-k} \underbrace{\Big(\gamma + (D-c'') - \lambda(T_i^* + c_i^*)\Big)}_{-\Delta(c'')} f(c'') - \lambda \int_{\underline{c}}^{c''} f(c)dc = 0 \tag{A.49}$$

$$\Rightarrow \Delta(c'')f(c'') = (k-1)\lambda F(c'') \tag{A.50}$$

That is, the extensive margin for cutting firms is found at the point where marginal damages when a cutting firm moves (left-hand side) are set equal to the overcompensation paid to the remaining firms that cut (right-hand side), located left from c''. From the first-order conditions for T_i^*:

$$\frac{1}{1-k}\Big(\gamma + (D - c'') - \lambda(T_i^* + c_i^*)\Big)f(c'') - \lambda\int_{\underline{c}}^{c''} f(c)dc + \frac{1}{k}\underbrace{\Big(\gamma - \lambda T_i^*\Big)}_{-\Delta(c')}f(c') - \lambda\int_{c'}^{\overline{c}} f(c)dc = 0$$

The sum of the two first terms is equal to zero by equation (A.49). Using this, and solving for the integral, we can write the first-order condition for T_i^* as:

$$\Delta(c')f(c') = -k\lambda\Big(1 - F(c')\Big). \tag{A.51}$$

Again, the optimal external margin is set by the trade-off between damages when a marginal non-cutting firm moves (left-hand side) and the overcompensation to all the mass of non-cutting firms, located right of point c' (right-hand side). Noting that $\Delta(c') = \Delta(c'') - (D - c'' + \lambda c_i^*)$, we can plug (A.51) into (A.50):

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

Since $c_i^* > c''$, we have $D - c'' - \lambda c_i^* > D - (1+\lambda)c_i^*$, and

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

$$\Rightarrow$$

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

$$\Rightarrow$$

$$c_i^* > \frac{D}{1+\lambda} - \frac{\lambda}{1+\lambda}\frac{F(c_i^*)}{f(c_i^*)}, \tag{A.52}$$

where the last two steps use

$$\lambda\frac{F(c'')}{f(c'')} - \lambda k\Big(\frac{F(c')}{f(c')} - \frac{F(c'')}{f(c'')} + \frac{1}{f(c')}\Big) < \lambda\frac{F(c'')}{f(c'')} \leq \lambda\frac{F(c_i^*)}{f(c_i^*)}.$$

The first inequality follows from the term in brackets being non-negative, as $c' \geq c''$ and $F(c)/f(c)$ is increasing in c by Assumption 1. The second inequality follows from $c_i^* > c''$ and Assumption 1. By (A.52), it follows that the emission price is always above that in case (i) of the Proposition ($k > 0$). Q.E.D.

Proof of Proposition 4.

We show that for strong positive correlation ($k > 1$) it is optimal to set a price for moving firms, $0 < c_j^*$, taking the form stated in Proposition 4. Then, we show that for weak positive correlation ($0 < k < 1$), it is optimal to set $0 < c_j^*$ under the condition stated in Proposition 4. Finally, we show that for negative ($k < 0$) correlation, price c_j^* cannot exceed the domestic price rendering the global market redundant (because moving firms have $c > c_j^*$).

Global market: $k > 1$**.** We proceed constructively and conjecture that $0 < c_j^* < c' < c_i^*$, so that firms with $c \leq c'$ move and firms with $c \leq c_j^*$ not only move but also cut abroad. Of the staying firms, $c > c'$, firms with $c \leq c_i^*$ cut. Consider the indifference at the extensive margin, for the marginal moving type $c = c'$:

$$C_i(c) - C_j(c) = b + kc \Leftrightarrow c - (T_i^* + c_i^*) = b + kc$$
$$\Leftrightarrow c = c' = \frac{(T_i^* + c_i^* + b)}{1-k}.$$

Type c' compares: stay and cut, or move and no cuts (because $c_j^* < c'$). With this definition of c', consider the welfare from choosing triple (c_i^*, c_j^*, T_i^*),

$$W_i = \int_{c'}^{c_i^*} \Big(\gamma + D - c - \lambda(T_i^* + c_i^*)\Big) f(c)dc + \int_{c_i^*}^{\bar{c}} \Big(\gamma - \lambda T_i^*\Big) f(c)dc + \underbrace{\int_{\underline{c}}^{c_j^*} \Big(D - (1+\lambda)c_j^*\Big) f(c)dc.}_{=\text{buying location } j \text{ reductions}}$$

For T_i^* to be optimal,

$$\frac{\partial W_i}{\partial T_i^*} = 0$$
$$\Leftrightarrow$$
$$\frac{-1}{1+k}\Big(\gamma + D - c' - \lambda(T_i^* + c_i^*)\Big) f(c') - \int_{c'}^{\bar{c}} \lambda f(c)dc = 0. \tag{A.53}$$

Similarly, for c_i^*

$$\frac{\partial W_i}{\partial c_i^*} = 0$$
$$\Leftrightarrow$$
$$\frac{-1}{1+k}\Big(\gamma + D - c' - \lambda(T_i^* + c_i^*)\Big) f(c') + \Big(D - (1+\lambda)c_i^*\Big) f(c_i^*) + \int_{c'}^{c^*} \lambda f(c)dc = 0 \tag{A.54}$$

Conditions (A.53)-(A.54) are equivalent to those for case $k > 1$ in Proposition 3. Therefore: As long as $c_j^* < c'$, the global price has no effect on the domestic policies.

Consider next optimal c_j^*:

$$\frac{\partial W_i}{\partial c_j^*} = 0 \Leftrightarrow \Big(D - (1+\lambda)c_j^*\Big) f(c_j^*) - (1+\lambda)F(c_j^*) = 0,$$

which gives the condition in Proposition 4, defining unique optimal $c_j^* > 0$. Thus, the market for reductions in j always emerges for $k > 1$ as stated in the Proposition.

No global market: $k < 0$. With negative correlation, firms with $c \geq c'$ move and marginal cutting firm at home is below this threshold: $0 < c_i^* < c'$. Thus, if there is an effective global price for pollution reductions, it must be $c_j^* > c_i^*$. We show that this is a contradiction.

The indifference at the extensive margin, for type $c = c'$, is

$$C_i(c) - C_j(c) = b + kc \Leftrightarrow -T_i^* - (c - c_j^*) = b + kc$$
$$\Leftrightarrow c = c' = \frac{c_j^* - (T_i^* + b)}{1+k}.$$

Firm c' does not cut at home ($C_i = -T_i^*$) but it cuts abroad and earns rent $(c - c_j^*)$, conditional on moving. With this c', we can define welfare:

$$W_i = \int_{\underline{c}}^{c_i^*} \Big(\gamma + D - c - \lambda(T_i^* + c_i^*)\Big) f(c)dc + \int_{c_i^*}^{c'} \Big(\gamma - \lambda T_i^*\Big) f(c)dc + \int_{c'}^{c_j^*} \Big(D - (1+\lambda)c_j^*\Big) f(c)dc$$

For optimal T_i^*, we have

$$\frac{\partial W_i}{\partial T_i^*} = 0$$
$$\Leftrightarrow$$
$$\frac{-1}{1+k}\Big(\gamma - \lambda T_i^*\Big) f(c') + \frac{1}{1+k}\Big(D - (1+\lambda)c_j^*\Big) f(c') - \int_{\underline{c}}^{c'} \lambda f(c)dc = 0.$$

The impact of c_j^* on welfare is

$$\frac{\partial W_i}{\partial c_j^*} =$$
$$\frac{1}{1+k}\Big(\gamma - \lambda T_i^*\Big) f(c') + \frac{-1}{1+k}\Big(D - (1+\lambda)c_j^*\Big) f(c')$$
$$+\Big(D - (1+\lambda)c_j^*\Big) f(c_j^*)$$
$$-(1+\lambda)\Big(F(c_j^*) - F(c')\Big).$$

We observe that

$$\frac{\partial W_i}{\partial T_i^*} = 0$$

$$\Rightarrow$$

$$\frac{\partial W_i}{\partial c_j^*} = -\lambda F(c') + \Big(D - (1+\lambda)c_j^*\Big) f(c_j^*) - (1+\lambda)\Big(F(c_j^*) - F(c')\Big) = 0$$

$$\Rightarrow D - (1+\lambda)c_j^* - \lambda \frac{F(c_j^*)}{f(c_j^*)} + \frac{F(c') - F(c_j^*)}{f(c_j^*)} = 0$$

Use $F(c') > F(c_j^*) \Rightarrow \frac{F(c')-F(c_j^*)}{f(c_j^*)} < 0$ to write:

$$D - (1+\lambda)c_j^* - \lambda \frac{F(c_j^*)}{f(c_j^*)} > 0$$

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

But this is a contradiction to $c_j^* > c_i^*$. Thus, the global price cannot exceed the local price, meaning that the moving firms have a higher cost than the price the policy maker is willing to pay.

Global market: $0 < k < 1$**.** We now construct the global market for weak positive correlation. Recall from the proof of Proposition 3 that, without the global price, there are two extensive margins: firms stay and cut if $c \leq c''$ and they stay without cutting if $c' \leq c$ where $c'' < c'$. We first show that it is not possible to create a market for leaving firms $c'' < c < c'$ with a global price such that $c_j^* > c''$. Suppose that this was possible: Consider threshold $c = c''$ for firms who cut home and abroad. Firms leave if the cost of cutting home is larger than cutting abroad:

$$\underbrace{c - T_i^* - c_i^*}_{\text{Cut home}} - \underbrace{(c - c_j^*)}_{\text{Cut abroad}} > \underbrace{b + kc}_{\theta}.$$

Firms leave if:

$$c < -\frac{b + T_i^* + c_i^* - c_j^*}{k} \text{ for } k > 0$$

Then, for $0 < k < 1$ it is not possible to create threshold where (1) those below that $c \leq c''$ stay and cut, (2) and $c > c''$ leave and cut. This suggests that we may construct the global market as follows: (i) firms leave and cut if $c \leq c'''$, (ii) firms stay and cut if $c''' < c \leq c''$, (iii) firms leave and pollute if $c'' < c \leq c'$ and (iv) firms stay and pollute if $c > c'$. Let us now construct these extensive margins one by one.

Margin c'''. Firm c''' is indifferent between leaving and cutting (to receive c_j^*) and staying and cutting (to receive $T_i^* + c_i^*$), if:

$$C_i(c) - C_j(c) = b + kc \Leftrightarrow c - (c_i^* + T_i^*) - (c - c_j^*) = b + kc \Leftrightarrow c''' = \frac{c_j^* - b - (c_i^* + T_i^*)}{k}.$$

Note that c''' decreasing in c_i^* and T_i^*. Paying more at home means that *less* firms leave.

Margin c''. Firm c'' is indifferent between leaving and polluting, and staying and cutting (to receive $T_i^* + c_i^*$), if:

$$C_i(c) - C_j(c) = b + kc \Leftrightarrow c - (c_i^* + T_i^*) - 0 = b + kc \Leftrightarrow c'' = \frac{b + c_i^* + T_i^*}{1-k}.$$

Margin c'. Firm c' is indifferent between staying and leaving, without cutting in either case:

$$C_i(c) - C_j(c) = b + kc \Leftrightarrow -T_i^* = b + kc \Leftrightarrow c' = \frac{-(b + T_i^*)}{k}.$$

With the definitions for c''', c'', and c', we define welfare to be maximized by choice (c_i^*, c_j^*, T_i^*)

$$W_i = \int_{\underline{c}}^{c'''} \Big(D - (1+\lambda)c_j^*\Big) f(c)dc + \int_{c'''}^{c''} \Big(\gamma + (D - c) - \lambda(T_i^* + c_i^*)\Big) f(c)dc + \int_{c'}^{\overline{c}} \Big(\gamma - \lambda T_i^*\Big) f(c)dc$$

The first-order condition with respect to c_j^* writes as:

$$\frac{1}{k}\Big(D - (1+\lambda)c_j^*\Big) f(c''') - \frac{1}{k}\Big(\gamma + D - c''' - \lambda(T_i^* + c_i^*)\Big) f(c''') - (1+\lambda)F(c''') = 0$$
$$\Rightarrow -\frac{1}{k}\underbrace{\Big(\gamma + (1+\lambda)c_j^* - c''' - \lambda(T_i^* + c_i^*)\Big)}_{=A} f(c''') - (1+\lambda)F(c''') = 0 \tag{A.55}$$

The first-order condition with respect to c_i^* is

$$\frac{1}{k}\Big(\gamma + (1+\lambda)c_j^* - c''' - \lambda(T_i^* + c_i^*)\Big) f(c''')$$
$$+\frac{1}{1-k}\Big(\gamma + (D - c'') - \lambda(T_i^* + c_i^*)\Big) f(c'') - \lambda\Big(F(c'') - F(c''')\Big) = 0 \tag{A.56}$$

Finally, the first-order condition with respect to T_i^* is

$$\frac{1}{k}\Big(\gamma + (1+\lambda)c_j^* - c''' - \lambda(T_i^* + c_i^*)\Big) f(c''')$$
$$+\frac{1}{1-k}\Big(\gamma + (D - c'') - \lambda(T_i^* + c_i^*)\Big) f(c'')$$
$$+\frac{1}{k}\Big(\gamma - \lambda T_i^*\Big) f(c') - \lambda(1 - F(c''')) = 0. \tag{A.57}$$

Combine (A.56) and (A.57),

$$\frac{1}{k}\Big(\gamma - \lambda T_i^*\Big) f(c') - \lambda(1 - F(c'')) = 0 \Rightarrow \Big(\gamma - \lambda T_i^*\Big) > 0$$

Evaluate conditions (A.56) and (A.57) at $c_j^* = c''' = 0$: they collapse to the ones from the local mechanism (proof of Proposition 3, case $0 < k < 1$). We have $A = \gamma - \lambda(T_i^* + c_i^*)$ in (A.55) for $c_j^* = c''' = 0$. Thus, if $A = \gamma - \lambda(T_i^* + c_i^*) < 0$, the marginal value c_j^* on W_i is strictly positive by the left-hand side of (A.55). This shows that a global price $c_j^* > 0$ must emerge, if the condition holds for the local mechanism. Q.E.D.

Proof of Proposition 5.

The Proposition considers policies abroad that are less ambitious than those at home: $c_j^* < c_i^*$. Given such c_j^* in place, welfare function in i is:

$$\max_{c_i^*, T_i^*} W_i = \int_{\underline{c}}^{c_i^*} \left(\gamma + D - c - \lambda(T_i^* + c_i^*)\right)\phi_i(C(c), c)dc + \int_{c_i^*}^{\overline{c}} \left(\gamma - \lambda T_i^*\right)\phi_i(C(c), c)dc + DQ_j(c_i^*, T_i^*)$$

where $Q_j(c_i^*, T_i^*)$ are the total reductions in country j. Firms' costs, determining their relocation decisions, are: $C(c) = c_j^*(c_i^*, T_i^*) - c_i^* - T_i^*$ if $c \leq c_j^*$; $C(c) = c - c_i^* - T_i^*$ if $c_j^* < c \leq c_i^*$; and $C(c) = -T_i^*$ if $c > c_i^*$. Here we allow both $c_j^*(\cdot)$ and $Q_j(\cdot)$ to depend on country i's policies: these interactions are determined by the choice of the instrument chosen by country j. We characterize (c_i^*, T_i^*) for the two policies, and then develop expression $Z(c)$ for comparing the two situations.

Price-based policy in j. For a price-based policy in country j, c_j^* is fixed and only the quantity $Q_j(c_i^*, T_i^*)$ depends on policies in i:

$$Q_j(c_i^*, T_i^*) = \int_{\underline{c}}^{c_j^*} \phi_j(C(c), c)dc = G(c_j^* - c_i^* - T_i^*)F(c_j^*)$$

$$\Rightarrow \frac{\partial Q_j(c_i^*, T_i^*)}{\partial c_i^*} = \frac{\partial Q_j(c_i^*, T_i^*)}{\partial T_i^*} = g(c_j^* - c_i^* - T_i^*)F(c_j^*). \tag{A.58}$$

Using equation (A.58), and denoting $-\Delta_G^{tax}(c) = \gamma - \lambda T_i^* - c - \lambda c_i^*$ and $-\Delta_D(c) = \gamma - \lambda T_i^* + D - c - \lambda c_i^*$, the first-order condition with respect to the base compensation level T_i^* is:

$$\begin{aligned} &\Big([\gamma - \lambda T_i^*]g(c_i^*) - \lambda(1 - G(c_i^*))\Big)\left(1 - F(c_i^*)\right) + \\ &\int_{\underline{c}}^{c_j^*} -\Delta_G^{tax}(c)g(c_j^*)f(c)dc - \lambda(1 - G(c_j^*))F(c_j^*) + \\ &\int_{c_j^*}^{c_i^*} -\Delta_L(c)g(c)f(c) - \lambda(1 - G(c))f(c)dc = 0. \end{aligned} \tag{A.59}$$

Similarly, local price c_i^* satisfies

$$\begin{aligned} &D - (1 + \lambda)c_i^* + \\ &\int_{\underline{c}}^{c_j^*} -\Delta_G^{tax}(c)\frac{g(c_j^*)}{1 - G(c_i^*)}\frac{f(c)}{f(c_i^*)}dc - \lambda\frac{1 - G(c_j^*)}{1 - G(c_i^*)}\frac{F(c_j^*)}{f(c_i^*)} + \\ &\int_{c_j^*}^{c_i^*} -\Delta_L(c)\frac{g(c)}{1 - G(c_i^*)}\frac{f(c)}{f(c_i^*)} - \lambda\frac{1 - G(c)}{1 - G(c_i^*)}\frac{f(c)}{f(c_i^*)}dc = 0 \end{aligned} \tag{A.60}$$

Here, the first-order conditions (A.59) and (A.60) characterize a best response to any foreign price that is equal to or below the local price. A comparison to the first order condition without policies in j in (10) reveals that introducing a tax in j has two countervailing effects. First, the term $-\Delta_G^{tax}(c)$ replaces the term $-\Delta_G(c) = \gamma - \lambda T_i^* + D - c - c_i^*$, with $-\Delta_G^{tax} < -\Delta_G$. This effect alone tends to decrease the optimal local price c_i^*. Second, for the first integral, the relocation cost function becomes $C(c_j^*) = c_j^* - c_i^* - T_i^*$ instead of $C(c) = c - c_i^* - T_i^*$ in (10). Combined with the hazard rate assumption (Assumption 1), we have $g(c_j^*)/(1-G(c_j^*)) > g(c)/(1-G(c))$ and so this effect tends to increase the optimal local price c_i^*.

Quantity-based policy in j. For a quantity-based policy in j, the externalities producer abroad are capped at Q_j^* (fixed), but the foreign externality price becomes endogenous:

$$Q_j = \int_{\underline{c}}^{c_j^*(c_i^*,T_i^*)} \phi_j(C(c),c)dc = G(c_j^*(c_i^*,T_i^*) - c_i^* - T_i^*)F(c_j^*(c_i^*,T_i^*))$$

Differentiate both sides with respect to c_i^*:

$$0 = -g(c_j^*)F(c_j) + \frac{\partial c_j^*}{\partial c_i^*}g(c_j^*)F(c_j^*) = \frac{\partial c_j^*}{\partial c_i^*}G(c_j^*)f(c_j^*)$$

$$\Rightarrow$$

$$\frac{\partial c_j^*(c_i^*,T_i^*)}{\partial c_i^*} = \frac{\partial c_j^*(c_i^*,T_i^*)}{\partial T_i^*} = \frac{g(c_j^*)F(c_j^*)}{g(c_j^*)F(c_j^*) + G(c_j^*)f(c_j^*)} = \xi \in (0,1). \tag{A.61}$$

where, as a shorthand, we have suppressed the arguments $g(c_j^*(c_i^*,T_i^*) - c_i^* - T_i^*) = g(c_j^*)$ and similarly for $G(c_j^*)$. Using equation (A.61), and denoting $-\Delta_G^{cap}(c) = (1-\xi)(\gamma - \lambda T_i^* + D - c - \lambda c_i^*)$, the first-order condition with respect to the base compensation level T_i^* is:

$$\begin{aligned}&\Big([\gamma - \lambda T_i^*]g(c_i^*) - \lambda(1-G(c_i^*))\Big)\left(1 - F(c_i^*)\right) + \\ &\int_{\underline{c}}^{c_j^*} -\Delta_G^{cap}(c)g(c_j^*)f(c)dc - \lambda(1-G(c_j^*))F(c_j^*) + \\ &\int_{c_j^*}^{c_i^*} -\Delta_L(c)g(c)f(c) - \lambda(1-G(c))f(c)dc = 0\end{aligned} \tag{A.62}$$

The first-order condition with respect to the local price c_i^* can be written as:

$$\begin{aligned}&D - (1+\lambda)c_i^* + \\ &\int_{\underline{c}}^{c_j^*} -\Delta_G^{cap}(c)\frac{g(c_j^*)}{1-G(c_i^*)}\frac{f(c)}{f(c_i^*)}dc - \lambda\frac{1-G(c_j^*)}{1-G(c_i^*)}\frac{F(c_j^*)}{f(c_i^*)} + \\ &\int_{c_j^*}^{c_i^*} -\Delta_L(c)\frac{g(c)}{1-G(c_i^*)}\frac{f(c)}{f(c_i^*)} - \lambda\frac{1-G(c)}{1-G(c_i^*)}\frac{f(c)}{f(c_i^*)}dc = 0\end{aligned} \tag{A.63}$$

Again, by comparing (10) to (A.63) we find two countervailing effects; first $-\Delta_G^{cap}(c) < -\Delta(c)$, and second, $C(c_j^*) = c_j^* - c_i^* - T_i^*$ replaces $C(c) = c - c_i^* - T_i^*$ in the first integral.

Comparison: expression $Z(c)$**.** If local price c_i^* is higher for a price-based scheme or a quantity based scheme boils down to comparing the first-order conditions (A.60) and (A.63). Recall that, for the sake of comparison, the two policies in j implement the same c_j^*. Then, a price-based policy in j leads to a higher local emission price c_i^* iff:

$$\int_{\underline{c}}^{c_j^*} -\Delta_G^{tax}(c)f(c)dc > \int_{\underline{c}}^{c_j^*} -\Delta_G^{cap}(c)f(c)dc$$

This holds true if $-\Delta_G^{tax} > -\Delta_G^{cap}$ for all $c \leq c_j^*$. Plug in Δ_G^{tax} and Δ_G^{cap}:

$$[\gamma - \lambda T_i^* - c - \lambda c_i^*] - (1-\xi)[\gamma - \lambda T_i^* + D - c - \lambda c_i^*]dc > 0.$$

Use the definition of ξ from (A.61):

$$[\gamma - \lambda T_i^* - c - \lambda c_i^*] - \frac{G(c_j^*)f(c_j^*)}{g(c_j^*)F(c_j^*) + G(c_j^*)f(c_j^*)}[\gamma - \lambda T_i^* + D - c - \lambda c_i^*] > 0$$

$$\Rightarrow$$

$$-\Delta(c)g(c_j^*)F(c_j^*) > DG(c_j^*)f(c_j^*)$$

$$\Rightarrow$$

$$Z(c, c_j^*) < 0$$

where $-\Delta(c) = \gamma - \lambda T_i^* - c - \lambda c_i^*$ is the marginal welfare effect of relocation when a firm cuts in both countries. Q.E.D.

General function forms

Derivation of equations (18) and (19). Unlike in the main section, the optimal solution no longer takes a bang-bang form (Lemmas 3 and 4 no longer hold) when the abatement cost function is not linear. The objective function writes as:

$$\max_{\substack{X_i(c),C_i(c),\\ X_j(c),C_j(c)}} W_i = \int_{\underline{c}}^{\overline{c}} \Big(\gamma + DX_i(c) - \lambda A(X_i(c),c) + \lambda C_i(c)\Big)\phi_i(C(c),c) + \\ \Big(DX_j(c) - (1+\lambda)A(X_j(c),c) + (1+\lambda)C_j(c)\Big)\phi_j(C(c),c)dc$$

so that the voluntary participation constraint $C_j(c) \leq 0$ holds, as well as the incentive compatibility, which can be written as $C_k'(c) = A_c(X_k(c),c)$, $k = i,j$ (proof is standard and is omitted, see for instance Baron and Myerson (1982)). Here, we focus on the cases where full separation is optimal, that is, where the non-monotonicity condition for neither $X_i(c)$ nor $X_j(c)$ binds. Hamiltonian for the problem is:

$$\mathcal{H} = \Big(\gamma + DX_i(c) - (1+\lambda)A(X_i(c),c) + \lambda C_i(c)\Big)\phi_i(C(c),c) + \Big(DX_j(c) - \\ (1+\lambda)A(X_j(c),c) + (1+\lambda)C_j(c)\Big)\phi_j(C(c),c) - \mu_i(c)A_c(X_i(c),c) - \mu_j(c)A_c(X_j(c),c) \tag{A.64}$$

where $\mu_k(c)$, $k = i,j$ denotes the co-state variables of the two incentive compatibility constraints. We assume that $X_i(c)$ and $X_j(c)$ are differentiable. Using Pontryagin's principle, the necessary conditions for the optimum are:

$$\Big(D - (1+\lambda)A_x(X_i(c),c)\Big)\phi_i(C(c),c) - \mu_i(c)A_{xc}(X_i(c),c) = 0 \tag{A.65}$$

$$\Big(D - (1+\lambda)A_x(X_j(c),c)\Big)\phi_j(C(c),c) - \mu_j(c)A_{xc}(X_j(c),c) = 0 \tag{A.66}$$

$$\mu_i'(c) = \Delta(c)\phi_i'(C(c),c) - \lambda\phi_i(C(c),c) \tag{A.67}$$

$$\mu_j'(c) = \Delta(c)\phi_j'(C(c),c) - (1+\lambda)\phi_j(C(c),c) \tag{A.68}$$

$$\mu_i(\underline{c}) = 0, \qquad \mu_j(\underline{c}) = 0 \tag{A.69}$$

Here $-\Delta(c) = \gamma + D(X_i(c) - X_j(c)) - C_i(c) - (1+\lambda)(T_i(c) - T_j(c))$ denotes the net losses from relocation. From (A.65) and (A.66) we can solve:

$$A_x(X_i(c),c) = \frac{D}{1+\lambda} + \frac{\mu_i(c)}{(1+\lambda)\phi_i(C(c),c)}A_{xc}(X_i(c),c) \tag{A.70}$$

$$A_x(X_j(c),c) = \frac{D}{1+\lambda} + \frac{\mu_j(c)}{(1+\lambda)\phi_j(C(c),c)}A_{xc}(X_j(c),c) \tag{A.71}$$

Integrating over (A.67) and (A.68), and fixing the lower bound by using the transversality conditions (A.69), we get:

$$\mu_i(c) = \int_{\underline{c}}^{c} \Big(\Delta(\tilde{c})\phi_i'(C(\tilde{c}),\tilde{c}) - \lambda\phi_i(C(\tilde{c}),\tilde{c})\Big)d\tilde{c} \tag{A.72}$$

$$\mu_j(c) = \int_{\underline{c}}^{c} \Big(\Delta(\tilde{c})\phi_j'(C(\tilde{c}),\tilde{c}) - (1+\lambda)\phi_j(C(\tilde{c}),\tilde{c})\Big)d\tilde{c} \tag{A.73}$$