EPRG WORKING PAPER

UNIVERSITY OF CAMBRIDGE | Energy Policy Research Group

Understanding overlapping policies: Internal carbon leakage and the punctured waterbed

EPRG Working Paper 1910

Cambridge Working Paper in Economics 1920

Grischa Perino, Robert A. Ritz and Arthur A. van Benthem

Abstract We present an integrated framework to understand the emissions impact of unilateral overlapping policies within a carbon-pricing system. "Internal carbon leakage" captures emissions displacement within the system (e.g., due to greater product imports from a neighbouring country). The waterbed effect captures the policy's interaction with the system's overall emissions cap. Current market rules in the reformed EU ETS, California's carbon market and RGGI feature "punctured" waterbeds that allow overlapping policies to affect aggregate emissions. We present simple formulae to estimate internal carbon leakage for different types of policy such as a carbon price floor (perhaps with a border tax adjustment), an energy efficiency program, and renewables support. The sign and magnitude of the climate benefit from an overlapping policy varies widely depending on its design, location and timing. Punctured waterbeds raise the stakes: well-designed overlapping policies can be much more climate-effective but others now backfire.

Keywords Cap-and-trade, carbon leakage, carbon price floor, carbon pricing, EU ETS, overlapping policy, hybrid policy, waterbed effect.

JEL Classification H23 (externalities), Q54 (climate)

Contact Robert Ritz – rar36@cam.ac.uk

Publication March 2019

Understanding overlapping policies: Internal carbon leakage and the punctured waterbed*

Grischa Perino[1] Robert A. Ritz[2] Arthur A. van Benthem[3,4]

March 4, 2019

Abstract

We present an integrated framework to understand the emissions impact of unilateral overlapping policies within a carbon-pricing system. "Internal carbon leakage" captures emissions displacement *within* the system (e.g., due to greater product imports from a neighbouring country). The waterbed effect captures the policy's interaction with the system's overall emissions cap. Current market rules in the reformed EU ETS, California's carbon market and RGGI feature "punctured" waterbeds that allow overlapping policies to affect aggregate emissions. We present simple formulae to estimate internal carbon leakage for different types of policy such as a carbon price floor (perhaps with a border tax adjustment), an energy efficiency program, and renewables support. The sign and magnitude of the climate benefit from an overlapping policy varies widely depending on its design, location and timing. Punctured waterbeds raise the stakes: well-designed overlapping policies can be much more climate-effective but others now backfire.

Keywords: Cap-and-trade, carbon leakage, carbon price floor, carbon pricing, EU ETS, overlapping policy, hybrid policy, waterbed effect.

JEL codes: H23 (externalities), Q54 (climate)

---

*We thank Severin Borenstein, Dallas Burtraw, Jim Bushnell, Sebastian Rausch and Herman Vollebergh for helpful comments and suggestions.

[1]Department of Socioeconomics, University of Hamburg, Germany. Email: Grischa.Perino@uni-hamburg.de. [2]Energy Policy Research Group (EPRG) and Judge Business School, University of Cambridge, United Kingdom. Email: rar36@cam.ac.uk. [3]The Wharton School, University of Pennsylvania, 1354 Steinberg Hall-Dietrich Hall, 3620 Locust Walk, Philadelphia, PA 19104, United States. Phone: +1 (215) 898-3013. Fax: +1 (215) 898-7635. Email: arthurv@wharton.upenn.edu. [4]National Bureau of Economic Research.

1 Introduction

Following reforms finally agreed in 2018, the European Union's emissions trading scheme (EU ETS) has been augmented with a Market Stability Reserve (MSR). The MSR's core feature is that, from 2023 onwards, it will cancel "excess" allowances (EUAs)—and thereby make the EU ETS's long-run emissions cap a function of market outcomes. This transforms a "plain vanilla" cap-and-trade design with a fixed cap into a complex variant of a "hybrid" policy instrument (Roberts and Spence, 1976; Pizer, 2002).

At the same time, Europe is seeing increasing unilateral action by individual EU member states wishing to "do more" than what the ETS centrally provides. For example, Great Britain has since 2013 imposed an additional Carbon Price Support on electricity generation to "top up" the EUA price; in December 2018, the Netherlands committed to introducing a similar policy.[1] Other examples include a plethora of support mechanisms for renewables and energy efficiency. These share a common feature: they are policies by an individual country aimed at an individual sector within a multi-country multi-sector ETS.

What is the climate benefit of such overlapping policies? Pre-MSR, the answer was clear. With a binding EU-wide emissions cap, any unilateral emissions reduction is exactly offset by an emissions increase elsewhere: the "waterbed effect" is 100% (Fankhauser et al., 2010; Böhringer, 2014; Edenhofer et al., 2018). The MSR, by canceling a fraction of surplus EUAs, punctures this waterbed. Recent estimates suggest that near-term unilateral action that reduces EU-wide emissions demand by 1 ton of CO_2 will, over time, translate into an emissions reduction of .5 tCO_2 or more (e.g., Perino, 2018). This enables unilateral action to have a global climate benefit.

Yet the crucial missing link in the argument lies in figuring out how large a unilateral action is actually required to achieve this 1 tCO_2 reduction in EU-wide emissions demand. The missing link is what we call "internal carbon leakage" within the EU ETS. Given the degree of market integration across Europe, a unilateral policy that reduces the domestic emissions of an individual country will often have knock-on effects on its neighbours. For example, a unilateral carbon price floor on Dutch electricity may lead to an increase in emissions—and higher allowance demand—from imports of German coal-fired power.

In this paper, we aim to fill this gap in the literature by providing a simple new integrated framework to understand the climate impacts of such unilateral action. We characterize internal leakage for three types of policy: cost-raising (e.g., a carbon price floor), demand-reducing (e.g., an energy efficiency program) and supply-increasing (e.g.,

[1]Unilateral action has been driven by concerns about low/volatile EUA prices and other market failures such as innovation externalities (Newbery et al., 2019). See also Flachsland et al. (2018) on the potential for an EU-wide carbon price floor.

renewables support). Our empirical illustrations show how our approach also speaks to overlapping policies within a wider carbon *tax* system and to cap-and-trade systems in North America. We hope our analysis will be of value to policymakers trying "in real time" to gauge the attractiveness of domestic climate initiatives.

Internal carbon leakage differs from the forms of "external" leakage typically considered in the literature. This includes, most prominently, leakage to foreign jurisdictions (Fowlie, 2009; Caron et al., 2015), e.g., leakage from the EU ETS to non-EU jurisdictions. It also includes leakage to sectors in the same jurisdiction that are not covered by the carbon-pricing system (Baylis et al., 2013), e.g., leakage from the EU ETS to uncovered sectors such as transport.

2 Conceptual framework

Consider unilateral action by an individual country i within a multi-country carbon-pricing system. Suppose that at a particular time t, holding the carbon price path $\boldsymbol{\tau}_t = (\tau_t, \tau_{t+1}, ..., \tau_T)$ fixed, this unilateral policy is successful at reducing country i's domestic demand for emissions, $\Delta e_{it} < 0$.

What is the policy's equilibrium impact on overall emissions, Δe_t^*? We answer this question using the following expression:

$$\Delta e_t^* = [1 - W_t][1 - L_{it}]\Delta e_{it}, \tag{1}$$

where L_{it} is the rate of internal carbon leakage associated with i's policy and W_t is the extent of the waterbed effect. Following IPCC (2007), we use the definition $L_{it} \equiv -\left[\Delta e_{-it}/\Delta e_{it}\right]_{\boldsymbol{\tau}_t \text{ fixed}}$, where Δe_{-it} is the change in the emissions demand of all other countries induced by i's policy—holding current and future carbon prices $\boldsymbol{\tau}_t$ fixed. Therefore $\Delta e_t \equiv [1 - L_{it}]\Delta e_{it}$ represents the net EU-wide change in allowance demand. This quantity will generally vary from one unilateral policy to another.

Given this, the waterbed effect $W_t \equiv 1 - \left[\Delta e_t^*/\Delta e_t\right]_{\boldsymbol{\tau}_t \text{ equilibrium}}$ captures that i's policy affects the system-wide carbon price path $\boldsymbol{\tau}_t$. This translates the net EU-wide change in emissions demand Δe_t into an equilibrium change in overall emissions Δe_t^*. The waterbed effect may be time-dependent but does not depend on the details of i's policy.

Under a "plain vanilla" cap-and-trade system with a fixed cap the waterbed effect $W_t = 1$, and so i's policy cannot shift overall emissions, $\Delta e_t^* = 0$. This conclusion applies *regardless* of the size of L_{it}.

The EU ETS's new MSR punctures the waterbed, $W_t \in (0, 1)$, so internal leakage becomes critical for the magnitude of Δe_t^*. We distinguish between three cases. First, with intermediate leakage, $L_{it} \in (0, 1)$, i's policy induces an overall reduction in emissions,

$\Delta e_t^* < 0$. Second, with negative leakage, $L_{it} < 0$, the same conclusion applies. It is strengthened whenever leakage is sufficiently negative; if $L_{it} < [-W_t/(1 - W_t)] \equiv \underline{L}_t < 0$, then the combination of leakage and waterbed induces a *stronger* overall emissions cut, $\Delta e_t^* < \Delta e_{it}$. Third, with leakage above 100%, $L_{it} \geq 1$, i's policy now backfires in that it leads to *higher* overall emissions, $\Delta e_t^* \geq 0 > \Delta e_{it}$. In short, the MSR raises the stakes: some policies can now backfire but well-designed policy can be much more climate-effective.

Unilateral action under a "plain vanilla" *carbon tax* system has a zero waterbed: there is no cap and no induced change in the carbon price. So this policy instrument is nested as $W_t = 0$ in our approach—and the same three leakage cases apply.

We do not attempt to quantify Δe_{it} for any particular policy but rather are interested in the mapping from a given Δe_{it} to Δe_t^*.

3 A model of internal carbon leakage

Next we present a new theory of internal carbon leakage arising from different types of unilateral overlapping policy. The model is static; time subscripts are omitted to simplify notation.

3.1 Model setup

There are two countries, i and j, where the latter can be interpreted as an aggregate of all countries except i. A representative firm in each country produces output x_k with a cost function $C_k(x_k)$, where $C_k'(\cdot), C_k''(\cdot) > 0$ for $k = i, j$. Firms face a common demand function $p(X)$, where $X \equiv x_i + x_j$ and $p'(\cdot) < 0$. Carbon emissions $e_k = \theta_k x_k$, where the emissions intensity $\theta_k > 0$ is a constant, so emissions are proportional to output. This is best thought of as the dirtiness of production at country k's marginal plant. All firms face a common carbon price τ.

Under perfect competition, the first-order condition for profit-maximization by the firm in country k is given by:

$$p(X) - C_k'(x_k) - \tau\theta_k = 0 \text{ for } k = i, j, \tag{2}$$

so price equals marginal cost where the latter includes production and carbon costs.

Some further definitions will be useful. First, let $\varepsilon^D \equiv -p(\cdot)/Xp'(\cdot) > 0$ be the price elasticity of demand. Second, let $\sigma_k \equiv x_k/X \in (0, 1)$ be the market share of country k's firm (so $\sigma_i + \sigma_j \equiv 1$). Third, let $\widehat{C}_k'(x_k) \equiv C_k'(x_k) + \tau\theta_k$ be k's *total* marginal cost and define $\eta_k^S \equiv x_k\widehat{C}_k''(x_k)/\widehat{C}_k'(x_k) > 0$ as its elasticity, also noting that $\widehat{C}_k''(x_k) \equiv C_k''(x_k)$. By

k's first-order condition, $x_k'(p) = 1/C_k''(x_k) > 0$, i.e., its supply curve is upward-sloping. So $\varepsilon_k^S \equiv p x_k'(p)/x_k(p) > 0$ is k's price elasticity of supply and, at the firm's optimum, $\eta_k^S = 1/\varepsilon_k^S$.

To simplify the analysis, we focus on *marginal* policies that perturb the equilibrium only by a small amount. Carbon leakage is thus defined as $L_i \equiv (-de_j/de_i)$. Given fixed emissions intensities, this is equivalent to $L_i = (\theta_j/\theta_i)(-dx_j/dx_i)$, where the first term is firms' "relative dirtiness" and the second term is *output* leakage.

3.2 Cost-raising unilateral policy

We begin with a unilateral policy λ_i that imposes an additional carbon price only on i's firms. Formally, i's firms now face a carbon price $\tau_i = \tau_i(\tau, \lambda_i)$, where $\frac{d}{d\tau}\tau_i(\tau, \lambda_i), \frac{d}{d\lambda_i}\tau_i(\tau, \lambda_i) > 0$, while j's firms continue to face $\tau_j = \tau$. A leading example is a unilateral carbon price floor on electricity generation designed to "top up" the EU ETS price, $\tau_i = \tau + \lambda_i$; our setup is more general in that it allows the top-up to be non-uniform.

For concreteness, we can think of the firms' common demand curve $p(X)$ as the willingness-to-pay of domestic consumers in country i; they are served partly by domestic production from within i and partly by imports from j.

Such unilateral action leads to an asymmetric cost shock, inducing i's firms to cut production and emissions but thereby raising the "competitiveness" of rivals in j.

Proposition 1 *A cost-increasing unilateral policy λ_i by country i has internal carbon leakage to country j equal to:*

$$L_i^{cost} = \frac{\theta_j}{\theta_i} \frac{\sigma_j}{\left(\sigma_j + \varepsilon^D/\varepsilon_j^S\right)} > 0.$$

Carbon leakage exceeds 100% whenever θ_j/θ_i is sufficiently larger than unity.

Carbon leakage is always positive because the underlying output leakage is positive—i's firms lose market share to j's. Output leakage is always less than 100% as i's policy raises the market price, i.e., there is positive carbon cost pass-through. Yet carbon leakage *can* exceed 100% if j's firms are sufficiently dirtier.

Proposition 1 (proof and subsequent formal results in Appendix A) provides a simple formula to quantify the leakage rate in terms of "industry primitives". The comparative statics are intuitive: leakage is high when i's market share is small, demand is relatively inelastic, and j's firms are more supply-responsive (e.g., because they hold significant spare capacity) and dirtier.

This formalizes the rationale for a regional coalition within the EU introducing a carbon price floor (Newbery et al., 2019): a coalition combines greater market share than single-country action and contains leakage.

Finally, the formula for L_i^{cost} does not depend on the precise functional form of $\tau_i = \tau_i(\tau, \lambda_i)$; this matters for the *absolute* output effects $(dx_i/d\lambda_i, dx_j/d\lambda_i)$ but not for the *relative* output effect—which is what L_i^{cost} captures.

3.3 Demand-reducing and supply-increasing unilateral policies

We now turn to a unilateral policy ϕ_i that reduces consumer demand or increases low-emissions production. Formally, write the demand curve as $p(X;\phi_i)$ where $\frac{\partial}{\partial \phi_i} p(X;\phi_i) < 0$. Both firms continue to face the carbon price τ.

An example of a demand-reducing policy is an energy-efficiency program that reduces consumer demand in country i; write direct demand as $(1-\phi_i)D(p)$, so $p(X;\phi_i) = D^{-1}\left(X/(1-\phi_i)\right)$. An example of a supply-increasing policy is a renewables support mechanism that brings in additional zero-emissions generation, so $p(X;\phi_i) = p(X+\phi_i)$. In both examples, $\frac{\partial}{\partial \phi_i} p(X;\phi_i) < 0$ at an interior equilibrium.

Such policies reduce the residual demand of i's firms—but also that of j's firms. In this sense, the policy coverage is broader than for the cost-raising policy.

Proposition 2 *Demand-reducing and supply-increasing unilateral policies ϕ_i by country i have internal carbon leakage to country j equal to:*

$$L_i^{demand} = L_i^{supply} = -\frac{\theta_j}{\theta_i} \frac{\sigma_j}{(1-\sigma_j)} \frac{\varepsilon_j^S}{\varepsilon_i^S} < 0.$$

Carbon leakage is negative: j's firms are directly affected and respond by cutting output and emissions. Akin to Proposition 1, leakage is more strongly negative where j's firms are dirtier, more supply-responsive and have greater market share. In addition, it is more pronounced if i's own supply-responsiveness is weaker; then i's output contraction is smaller relative to j's.

The leakage rate does not depend on *any* demand characteristics, including the form of $p(X;\phi_i)$ and ε^D. The reason stems from policy coverage: different demand-side impacts, in general, affect i's and j's *absolute* production responses but, because they are hit by an *identical* demand shock, the *relative* magnitude does not depend on any demand parameters. At the margin, both policies have identical carbon leakage: $L_i^{\text{demand}} = L_i^{\text{supply}}$.

3.4 Extensions

We extend the model with an "end-of-pipe" abatement technology such as carbon capture & storage that cleans up production *ex post*. A firm's investment incentive rises with its domestic carbon price. For a cost-raising policy, we show (Proposition 1A) that such abatement reduces i's leakage rate—though, as in Proposition 1, it remains positive and exceeds 100% for θ_j/θ_i sufficiently high. Proposition 2 is unaffected because carbon prices (τ_i, τ_j) remain the same and so i's policy induces no incremental abatement (Proposition 2A).

A cost-increasing policy such as a unilateral carbon price floor is sometimes accompanied by a border tax adjustment (BTA) on imports. This increases policy coverage: the higher carbon price now applies to both i's and j's firms. In a further extension, we show (Proposition 1B) that a BTA can induce j's firms to also cut emissions so that $L_i^{\text{cost+BTA}} < 0$; in the special case where i and j have identical emissions intensities $(\theta_i = \theta_j)$, leakage is exactly as in Proposition 2, $L_i^{\text{cost+BTA}} = L_i^{\text{demand}} = L_i^{\text{supply}} < 0$. All proofs are in Appendix B.

4 The punctured waterbed under the reformed EU ETS

We next derive the waterbed effect W_t for the reformed EU ETS. The new MSR works as follows. If the total number of allowances in circulation (the "bank") exceeds 833 million at the end of a given year (in 2017 or later), then the number of allowances auctioned in the 12 months following October of the following year (but not before January 2019) is reduced by a certain percentage of the size of the bank as shown in Table 1. Allowances withheld are placed in the MSR and released in installments of 100 million/year once the bank has dropped below 400 million. We label $t_{B=833}$ the year in which the bank drops below the 833 million threshold and the MSR hence stops taking in allowances.

Time at which bank exceeds 833 million allowances (on December 31st)	Intake rate (%)
2017	16*
2018 - 2021	24
2021 - $t_{B=833}$	12

Table 1: Intake rates for the EU ETS Market Stability Reserve (MSR)
* Two-thirds of 24 percent because the withdrawals that would be due in Oct.-Dec. 2018 do not materialize (European Commission, 2018).

Starting in 2023, the maximum number of allowances held in the MSR is limited to the number auctioned in the previous year. The target share of auctioning in Phase 4 is 57% (European Parliament and Council, 2018) with the remaining allowances being freely allocated. Allowances stored in the MSR in excess of this upper bound are permanently canceled. The long-run cap is thus a function of market outcomes—and can be impacted by overlapping policies within the EU ETS.

We compute W_t as follows. The net change in allowance demand Δe_t is assumed to translate into an instantaneous change in the number of banked allowances assuming a fixed carbon price path. With a fixed cap and no reserve mechanism, the bank would hence shift by Δe_t from that year forward. We call the associated change in allowances transferred into the MSR and canceled, the direct impact.

We abstract from indirect impacts. One is that Δe_t induces a small adjustment of the carbon price and emission paths and hence the evolution of the bank. These effects are small in the period just following t because the bank is a cumulative measure that exhibits inertia towards small flow adjustments. A further indirect effect is changes to the years in which the MSR thresholds are passed; given our focus on marginal policies, the associated error is small.

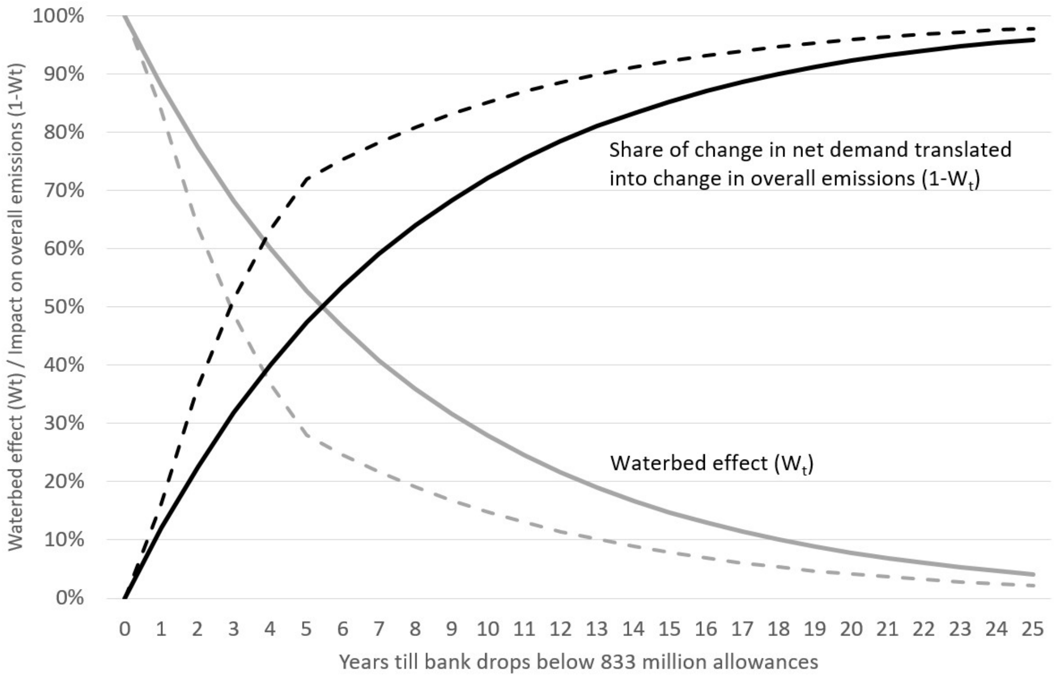
Everything else equal, adding one allowance to the bank triggers a sequence of transfers to the MSR. Because only a share ρ_t of the added amount is transferred in the first year, the remainder $(1-\rho_t)$ adds to the bank in the following year and again induces a transfer at rate ρ_{t+1}, i.e. $(1-\rho_t)\rho_{t+1}$, and so on. Using this rule, we compute the total direct waterbed effect as a function of the time t of unilateral action and $t_{B=833}$:

Lemma 1 *For unilateral action in year t and the bank dropping below 833 million allowances in year $t_{B=833}$, the (direct) waterbed effect under the MSR is given by:*

$$\begin{aligned} W_{t,t_{B=833}} &= (1-.16)^{\max[0,\min[2018,t_{B=833}]-\max[2017,t]]} \\ &\times (1-.24)^{\max[0,\min[2022,t_{B=833}]-\max[2018,t]]} \\ &\times (1-.12)^{\max[0,\max[2022,t_{B=833}]-\max[2022,t]]}. \end{aligned} \tag{3}$$

The share of a net change in allowance demand in year t that ends up being canceled is $1-W_{t,t_{B=833}}$. Figure 1 illustrates. The degree of puncture depends crucially on the number of years left until $t_{B=833}$. This date is subject to substantial uncertainty, with estimates ranging from 2020 (Perino, 2018) to the second half of the 2030s (Beck and Kruse-Andersen, 2018; Quemin and Trotignon, 2018), and $t_{B=833}=2030$ as a mid-range value (Vollebergh, 2018). The closer the change in net demand occurs to $t_{B=833}$, the smaller its impact on the long-run cap.

Figure 1: The evolving waterbed effect in the EU ETS

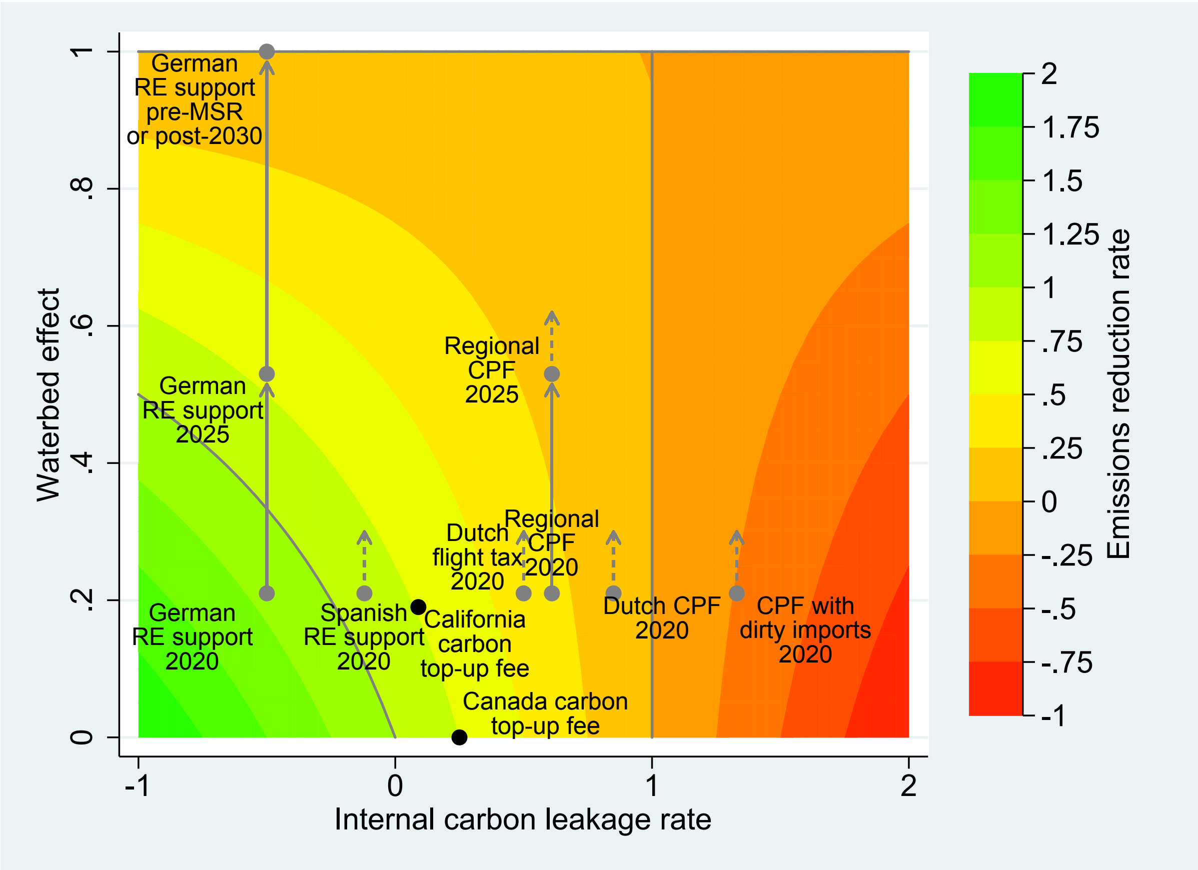


Notes: Eventual impact of a marginal change in net allowance demand on overall EU ETS emissions $1 - W_t$ (black) and the waterbed effect W_t (grey) as functions of the number of years until the aggregate bank drops below 833 million allowances $t_{B=833}$. Dashed: Effect in 2017 or earlier. Solid: Effect in 2022 or later. Effects for 2018-2021 in between (not shown). Calculations assume fixed carbon price path.

5 Illustrations of unilateral overlapping policies

We now illustrate how real-world overlapping policies fit into our model's framework. Our main outcome of interest is the emissions reduction rate $R_{it} = (1 - L_{it})(1 - W_t)$, the ratio of equilibrium system-wide to domestic emissions reductions. Figure 2 plots the contour lines of R_{it} in (L, W)-space along with various policy examples (described in more detail in Appendix C).

Figure 2: Unilateral policies facing internal carbon leakage and a waterbed effect



Notes: Figure shows the contour plot of the emissions reduction rate $R_{it} = (1 - L_{it})(1 - W_t)$ of various policies discussed in this section. Solid black lines indicate the contour lines where $R_{it} = 0$ (when $L = 1$ or $W = 1$) and $R_{it} = 1$ (bottom left). Dashed grey arrows indicate that, in the EU ETS, a policy's R_{it} moves towards zero as t approaches $t_{B=833}$ and $W_t \rightarrow 1$. We assume $t_{B=833} = 2030$. Solid grey arrows show specific shifts in time for the German renewable energy support schemes and for a proposed regional carbon price floor.

5.1 Overlapping policies in the EU ETS

As shown in Figure 1, the waterbed effect in the EU ETS depends on the year in which the policy takes effect t and the number of years until the bank drops below 833 million allowances $t_{B=833}$. We consider $t_{B=833} = 2030$ and contrast policies acting in years $t = 2020$, 2025 and 2030. As time moves on, W_t increases from 0.21 to 0.53 to 1 and all European policies in Figure 2 (in grey) move north, as indicated by the dotted lines. The internal leakage rate L_{it} is policy- and time-specific.

We first consider unilateral cost-raising policies such as a national carbon price floor (CPF) for electricity generation—as announced in 2018 by the Dutch Government. Proposition 1 shows that such policies, if binding, suffer from intra-EU leakage. We expect high leakage for small countries (high σ_j) that are strongly interconnected to neighbours with flexible yet dirty supply (high $\varepsilon_j^S, \theta_j/\theta_i$). Consistent with this, recent estimates find $L \simeq 0.85$ for the Dutch CPF, while a regional CPF including the Benelux, France and Germany faces $L = 0.61$ (Frontier Economics, 2018; Vollebergh, 2018). Such CPFs in small interconnected countries are unlikely to reduce EU-wide emissions by much, with $R = 0.15$ in 2020 even under the punctured waterbed (see Figure 2).[2] As more countries join the CPF, R rises to 0.30. Furthermore, the solid grey arrow shows that the regional CPF's R decreases to 0.18 by 2025, so early action is preferable.

Cost-raising policies can backfire if imports are substantially dirtier than domestic production (see Proposition 1). We plot a hypothetical "CPF with dirty imports" for which $L = 1.33$ such that EU-wide emissions *increase*, $R < 0$.[3] Since this policy lies to the right of the $R = 0$ contour line, the negative effect gets *weaker* over time as the waterbed effect gets stronger. Post-2030, all unilateral policies within the EU ETS end up at $R = 0$.

As another example of a cost-raising unilateral policy, several European countries, such as Austria, Germany, Norway and Sweden, have aviation taxes. The Netherlands had a short-lived aviation tax in 2008-9, which faced leakage of 50% as passengers substituted to Belgian and German airports (Gordijn and Kolkman, 2011). If reintroduced in 2020, as planned by the current Dutch government (Dutch News, 2018), we estimate $R = 0.38$. There is some broader evidence that aviation taxes are most likely where leakage is mitigated—e.g., in high-population countries (low σ_j) or countries far away from low-tax airports abroad (high ε_j^S) (PricewaterhouseCoopers, 2017).

Unilateral renewable support schemes fit our definition of supply-increasing policies. Abrell et al. (2019) estimate *negative* carbon leakage as zero-carbon energy offsets imported gas- and coal-fired electricity in Germany ($L = -0.50$) and Spain ($L = -0.12$). Similarly, a German government report finds $L = -0.65$ (Klobasa and Sensfuss, 2016). Negative leakage corresponds to Proposition 2 and should also occur for energy-efficiency policies.

[2] Internal carbon leakage may be weaker for Great Britain's carbon price floor due to less interconnection but we are not aware of any empirical estimates.

[3] We assume $\theta_j/\theta_i = 2$, $\varepsilon_j^S = 5 \Leftrightarrow \eta_j^S = 0.2$, $\sigma_j = 0.2$, and $\varepsilon^D = 0.5$.

5.2 Overlapping policies in North America

In a stochastic sense, several carbon-pricing systems in North America have a punctured waterbed. This happens in the presence of price floors and ceilings in allowance auctions, combined with uncertainty over when the system will trade in the "intermediate range" (full waterbed) vs. at the price floor/ceiling (zero waterbed).

California and Québec have a joint carbon market with an auction price floor and, in a recent proposal to take effect in 2021, a hard price ceiling (Politico, 2018). Unsold allowances during periods when the price floor binds—which it did in various auctions—are essentially retired. Borenstein et al. (2017) and Borenstein et al. (2018) report a post-reform estimate of 47% (34%) that the price floor (ceiling) binds. Thus, in expected terms, $W = 1 - 0.47 - 0.34 = 0.19$.

Now consider a counterfactual Western Climate Initiative (WCI) joined by states surrounding California. If California imposes an additional cost-raising policy such as a fixed carbon top-up fee, there is internal leakage to neighbouring states. Caron et al. (2015) provide a relevant leakage estimate of $L = 0.09$ assuming that—as the current market rules specify—there is a border tax adjustment and resource shuffling is prevented. Interestingly, they also report that internal carbon leakage for the electricity sector only is negative (-13%), consistent with our theory in Section 3.4. In Caron et al. (2015), this is more than offset by positive leakage to other sectors in the U.S. and internationally—yielding an aggregate leakage rate of 9%.

The Regional Greenhouse Gas Initiative (RGGI) that caps CO_2 emissions from electricity in ten Northeastern states is similar. Its CPF was binding for many years, but allowances now trade at higher prices. New York is currently considering an additional carbon fee, which would also apply to imported electricity from other RGGI states; Shawhan et al. (2018) find negative (output) leakage of $L = -0.77$.

Finally, Canada has a national minimum carbon tax of $20 per ton in 2019, increasing to $50 by 2022. Some provinces, such as Alberta and British Columbia, already had in place carbon taxes with a price above the federal tax. Such unilateral carbon taxes face no waterbed effect but may suffer from leakage to other provinces—though we are not aware of direct leakage estimates, Murray and Rivers (2015) and Yamazaki (2017) suggest that British Columbia's carbon tax has had negligible or modest effects on the aggregate economy, suggesting leakage is modest and so Figure 2 plots this policy assuming $L = 0.25$.

In sum, the empirical results are consistent with our theory of internal carbon leakage. Leakage for cost-raising policies (e.g., CPFs and flight taxes in Europe) is positive, except when imports are taxed (e.g., carbon fees in California and New York). Supply-increasing policies, such as German and Spanish renewables support, have negative leakage.

6 Conclusions

We have developed a simple new framework to understand unilateral overlapping policies within a wider carbon-pricing system. Design matters in that different policy types have very different leakage implications. Space matters as internal carbon leakage rates can differ substantially across industries and countries. Time matters in that it affects the magnitude of the waterbed. The EU ETS with the MSR and overlapping policies is about as complex as tackling local pollutants with highly heterogeneous marginal damages—and the uniform price rule is no longer straightforwardly appropriate.

On the empirical side, we illustrated how observed policies fit into our framework. Current market rules in the reformed EU ETS, California and RGGI feature punctured waterbeds that allow unilateral policies to affect aggregate emissions. Policies such as national energy conservation programs and renewable energy support schemes typically have *negative* internal carbon leakage. Yet there are surprisingly few estimates of internal leakage in the literature—and better information could substantially improve future policy-making.

We should acknowledge that our analysis is only a first cut. We have focused narrowly on emissions impacts—only one (important) component of a welfare analysis—and have not formally ranked different policy options. These are important topics for future research. Finally, we did not analyze "external" carbon leakage to outside jurisdictions; for policies affecting the electricity sector this is typically negligible.

Online Appendix A: Proofs of Propositions 1 and 2

Proof of Proposition 1. Given the asymmetric carbon prices, country k's first-order condition from (2) now becomes $p(X) - C_k'(x_k) - \tau_k\theta_k = 0$ for $k = i, j$, where $\tau_i = \tau_i(\tau, \lambda_i)$ and $\tau_j = \tau$. Write the resulting equilibrium outputs in terms of i's unilateral policy as $(x_i(\lambda_i), x_j(\lambda_i))$. What is the rate of output leakage? At equilibrium, j's first-order condition is:

$$p(x_i(\lambda_i) + x_j(\lambda_i)) - C_j'(x_j(\lambda_i)) - \tau\theta_j = 0.$$

Differentiating with respect to i's policy gives:

$$p'(x_i(\lambda_i) + x_j(\lambda_i)) \left(\frac{dx_i}{d\lambda_i} + \frac{dx_j}{d\lambda_i} \right) - C_j''(x_j(\lambda_i)) \frac{dx_j}{d\lambda_i} = 0,$$

where it is easy to check that:

$$\frac{dx_i}{d\lambda_i} = \frac{dx_i}{d\tau_i}\frac{d\tau_i}{d\lambda_i} < 0 \text{ and } \frac{dx_j}{d\lambda_i} = \frac{dx_j}{d\tau_i}\frac{d\tau_i}{d\lambda_i} > 0.$$

Rearranging yields a first expression for output leakage:

$$-\frac{dx_j/d\lambda_i}{dx_i/d\lambda_i} = \left[\frac{-p'(X)}{-p'(X) + C_j''(x_j)}\right]_{x_i = x_i(\lambda_i), x_j = x_j(\lambda_i)} \in (0,1).$$

This expression holds regardless of the precise functional form of $\tau_i(\tau, \eta_i)$; the only thing that matters is that $\frac{d}{d\lambda_i}\tau_i(\tau, \lambda_i) > 0$ leads to $dx_i/d\lambda_i < 0$. Now turn this into a more empirically-useful expression in two steps. First, using the definition of the price elasticity of demand $\varepsilon^D \equiv -p(\cdot)/Xp'(\cdot) > 0$, we have $-p'(X) = (1/\varepsilon^D)p(X)/X$. Second, using the definition of j's elasticity of total marginal cost $\eta_j^S \equiv x_j\widehat{C}_j''(x_j)/\widehat{C}_j'(x_j) > 0$, where $\widehat{C}_j'(x_j) \equiv C_j'(x_j) + \tau\theta_j = p(X)$ and $\widehat{C}_j''(x_j) \equiv C_j''(x_j)$, we have:

$$C_j''(x_j) = \frac{x_jC_j''(x_j)}{\widehat{C}_j'(x_j)}\frac{\widehat{C}_j'(x_j)}{x_j} = \frac{x_j\widehat{C}_j''(x_j)}{\widehat{C}_j'(x_j)}\frac{\widehat{C}_j'(x_j)}{x_j} = \eta_j^S\frac{p(X)}{X}\frac{1}{\sigma_j}, \tag{4}$$

where the last expression uses the definition of market share, $\sigma_j \equiv x_j/X$. Putting these two steps together gives another expression for output leakage:

$$-\frac{dx_j/d\lambda_i}{dx_i/d\lambda_i} = \frac{\sigma_j}{(\sigma_j + \varepsilon^D\eta_j^S)} > 0. \tag{5}$$

This converts into carbon leakage $L_i^{\text{cost}} = (\theta_j/\theta_i)\left[\sigma_j/(\sigma_j + \varepsilon^D/\varepsilon_j^S)\right] > 0$, as claimed, using the equilibrium relationship of price elasticity of j's supply curve, $\eta_j^S = 1/\varepsilon_j^S$, together with the relationship between output and carbon leakage, $L_i = (\theta_j/\theta_i)\,(-dx_j/dx_i)$.

Proof of Proposition 2. As in the proof of Proposition 1, write equilibrium outputs in terms of i's unilateral policy as $(x_i(\phi_i), x_j(\phi_i))$. At equilibrium, k's first-order condition from (2) becomes:

$$p(x_i(\phi_i) + x_j(\phi_i); \phi_i) - C_k'(x_k(\phi_i)) - \tau\theta_k = 0.$$

Differentiating this condition for i and j, respectively, yields:

$$\frac{\partial}{\partial\phi_i}p(x_i(\phi_i) + x_j(\phi_i); \phi_i) + p'(x_i(\phi_i) + x_j(\phi_i); \phi_i)\left(\frac{dx_i}{d\phi_i} + \frac{dx_j}{d\phi_i}\right) - C_i''(x_i(\phi_i))\frac{dx_i}{d\phi_i} = 0,$$
$$\frac{\partial}{\partial\phi_i}p(x_i(\phi_i) + x_j(\phi_i); \phi_i) + p'(x_i(\phi_i) + x_j(\phi_i); \phi_i)\left(\frac{dx_i}{d\phi_i} + \frac{dx_j}{d\phi_i}\right) - C_j''(x_j(\phi_i))\frac{dx_j}{d\phi_i} = 0.$$

This implies that, in equilibrium, output responses satisfy:

$$C_i''(x_i(\phi_i))\frac{dx_i}{d\phi_i} = C_j''(x_j(\phi_i))\frac{dx_j}{d\phi_i} < 0.$$

Some rearranging, using (4) from the proof of Proposition 1 to rewrite $C_i''(\cdot), C_j''(\cdot)$, gives an expression for output leakage:

$$-\frac{dx_j/d\lambda_i}{dx_i/d\lambda_i} = -\frac{\sigma_j}{(1-\sigma_j)}\frac{\varepsilon_j^S}{\varepsilon_i^S} < 0. \tag{6}$$

So carbon leakage $L_i = -(\theta_j/\theta_i)[\sigma_j/(1-\sigma_j)](\varepsilon_j^S/\varepsilon_i^S)$, as claimed, using the relationship $L_i = (\theta_j/\theta_i)(-dx_j/dx_i)$. Since this expression does not depend on the form of $\frac{\partial}{\partial \phi_i} p(x_i(\phi_i)$, we conclude that $L_i^{\text{demand}} = L_i^{\text{supply}}$.

Online Appendix B: Extensions to the model of internal carbon leakage

Leakage with an end-of-pipe abatement technology

Consider the same model setup as in Section 3.1 with the addition of investment in an end-of-pipe (EOP) abatement technology. Firm k now solves $\max_{x_k, e_k} \Pi_k = px_k - C_k(x_k) - \tau_k e_k - \phi_k(\theta_k x_k - e_k)$, where its abatement cost function satisfies $\phi_k'(\cdot), \phi_k''(\cdot) > 0$ if abatement investment takes place (i.e., $e_k < \theta_k x_k$) and $\phi_k(0) = \phi_k'(0) = 0$ if it does not (i.e., $e_k = \theta_k x_k$).

The two first-order conditions are now:

$$\partial \Pi_k / \partial x_k = p - C_k'(x_k) - \theta_k \phi_k'(\cdot) = 0 \text{ and } \partial \Pi_k / \partial e_k = -\tau_k + \phi_k' = 0, \tag{7}$$

where the former is a generalized version of price equals marginal cost and the latter says that the marginal abatement cost equals the carbon price. Note that k's abatement incentive rises with its domestic carbon price (since $\phi_k''(\cdot) > 0$).

The crucial point is that, taken together, these first-order conditions imply $p(X) - C_k'(x_k) - \tau_k \theta_k = 0$, *exactly* as in (2) for the benchmark model. In other words, the addition of an EOP technology does not affect the product-market outcome. Therefore, for all three types of unilateral policy, *output* leakage $(-dx_j/dx_i) > 0$ remains exactly as above.

A difference now does arise in terms of *carbon* leakage. For j's firm, it remains true that $de_j = \theta_j(dx_j)$, as in the benchmark model, again for all three policy types. So the remaining question is how different unilateral policies affect i's emissions $e_i = \theta_i x_i - \phi_i'^{-1}(\tau_i)$.

The following two results, as analogues to Propositions 1 and 2, characterize the rate of internal carbon leakage for our three types of policy.

Proposition 1A. *With an end-of-pipe (EOP) abatement technology, a cost-increasing unilateral policy λ_i by country i has internal carbon leakage to country j equal to:*

$$L_i^{\text{cost+EOP}} = \frac{\theta_j}{\theta_i} \frac{\dfrac{\sigma_j}{(\sigma_j + \varepsilon^D/\varepsilon_j^S)}}{\left[1 + \dfrac{d\phi_i'^{-1}(\tau_i(\tau,\lambda_i))/d\lambda_i}{-dx_i/d\lambda_i}\right]} < L_i^{\text{cost}}.$$

Carbon leakage exceeds 100% whenever θ_j/θ_i is sufficiently larger than unity.

Proof. By (7), country k's emissions are $e_k = \theta_k x_k - \phi_k'^{-1}(\tau_k)$. For country j, as its carbon price remains fixed at $\tau_j = \tau$, so the change in its emissions due to i's policy satisfies $de_j/d\lambda_i = \theta_j(dx_j/d\lambda_i)$. The rate of carbon leakage can be thus written as:

$$L_i \equiv -\frac{de_j/d\lambda_i}{de_i/d\lambda_i} = -\frac{\theta_j(dx_j/d\lambda_i)}{\dfrac{d}{d\lambda_i}\left[\theta_i x_i - \phi_i'^{-1}(\tau_i(\tau,\lambda_i))\right]} = \frac{\theta_j}{\theta_i}\frac{\left(-\dfrac{dx_j/d\lambda_i}{dx_i/d\lambda_i}\right)}{\left[1 + \dfrac{d\phi_i'^{-1}(\tau_i(\tau,\lambda_i))/d\lambda_i}{-dx_i/d\lambda_i}\right]}.$$

By the first-order conditions (7), end-of-pipe abatement does not affect the product-market outcome and thus also not the rate of output leakage—which remains as given by (5) in the proof of Proposition 1, $-(dx_j/d\lambda_i)/(dx_i/d\lambda_i) = \sigma_j/(\sigma_j + \varepsilon^D/\varepsilon_j^S) > 0$. This yields $L_i^{\text{cost+EOP}}$ as claimed, and $L_i^{\text{cost+EOP}} < L_i^{\text{cost}}$ holds because $dx_i/d\lambda_i < 0$ from the proof of Proposition 1 and

$$\frac{d}{d\lambda_i}\phi_i'^{-1}(\tau_i(\tau,\lambda_i)) = \frac{1}{\phi_i''(\phi_i'^{-1}(\tau_i(\tau,\lambda_i))}\frac{d}{d\lambda_i}\tau_i(\tau,\lambda_i) > 0$$

holds given the assumptions $\phi_i'(\cdot), \phi_i''(\cdot) > 0$ and $\frac{d}{d\lambda_i}\tau_i(\tau,\lambda_i) > 0$. ■

As is intuitive, greater abatement activity tends to mitigate the carbon leakage induced by a cost-raising policy: a given output reduction by i's firms now comes with a stronger reduction in emissions. However, it remains true that carbon leakage can be high, including above 100%, if j's firms are sufficiently dirtier than i's.

Proposition 2A. *With an end-of-pipe (EOP) abatement technology, demand-reducing and supply-increasing unilateral policies ϕ_i by country i have internal carbon leakage to country j equal to*:

$$L_i^{\text{demand+EOP}} = L_i^{\text{supply+EOP}} = -\frac{\theta_j}{\theta_i}\frac{\sigma_j}{(1-\sigma_j)}\frac{\varepsilon_j^S}{\varepsilon_i^S} = L_i^{\text{demand}} = L_i^{\text{supply}} < 0.$$

Proof. By (7), country k's emissions are $e_k = \theta_k x_k - \phi_k'^{-1}(\tau_k)$. For both countries, as the system-wide carbon price here remains fixed at $\tau_i = \tau_j = \tau$, the emissions changes due to i's policy satisfy $de_k/d\lambda_i = \theta_k(dx_k/d\lambda_i)$. By the first-order conditions (7), end-of-pipe abatement does not affect the product-market outcome and thus also not the rate of output leakage—which remains as given by (6) in the proof of Proposition 2, $-(dx_j/d\lambda_i)/(dx_i/d\lambda_i) = -[\sigma_j/(1-\sigma_j)](\varepsilon_j^S/\varepsilon_i^S) < 0$. So carbon leakage is exactly as in Proposition 2, $L_i^{\text{demand+EOP}} = L_i^{\text{supply+EOP}} = L_i^{\text{demand}} = L_i^{\text{supply}}$, as claimed. ■

For demand-reducing and supply-increasing policies, firm i's carbon price remains unchanged and so they do not induce any incremental abatement effect. Hence emissions changes are driven solely by output changes, and since end-of-pipe abatement does not affect output, carbon leakage remains exactly as in Proposition 2 in the main text.

Leakage of a cost-raising policy with a border tax adjustment

Consider the same model setup as in Section 3.1 and the same cost-raising policy as in Section 3.2, now with the addition of a border tax adjustment (BTA) that accompanies i's policy. This increases policy coverage: a higher carbon price now applies to *both* i's and j's firms. Again denoting the policy as λ_i, k's firms ($k = i, j$) now face a common carbon price $\tau_k = \tau_i(\tau, \lambda_i)$, strictly increasing in each argument $\frac{d}{d\tau}\tau_i(\tau, \lambda_i), \frac{d}{d\lambda_i}\tau_i(\tau, \lambda_i) > 0$.

We obtain the following result:

Proposition 1B. *With a border tax adjustment (BTA), assuming $\theta_i/\theta_j > \sigma_j/(\sigma_j + \varepsilon^D/\varepsilon_j^S) \equiv \underline{\theta} \in (0,1)$, a cost-increasing unilateral policy λ_i by country i reduces its domestic emissions and has internal carbon leakage to country j equal to:*

$$L_i^{\text{cost+BTA}} = -\frac{\theta_j}{\theta_i} \frac{\left[\left(\frac{\theta_j}{\theta_i} - 1\right) + \frac{\theta_j}{\theta_i}\frac{\varepsilon^D}{\varepsilon_i^S}\frac{1}{(1-\sigma_j)}\right]}{\left[\left(1 - \frac{\theta_j}{\theta_i}\right) + \frac{\varepsilon^D}{\varepsilon_j^S}\frac{1}{\sigma_j}\right]}.$$

Leakage is negative $L_i^{\text{cost+BTA}} < 0$ if and only if $\theta_i/\theta_j < [(1-\sigma_j) + \varepsilon^D/\varepsilon_i^S]/(1-\sigma_j) \equiv \overline{\theta}$, where $\overline{\theta} > 1$. If countries have identical emissions intensities with $\theta_i = \theta_j$, then leakage is identical to demand-reducing and supply-increasing policies, $L_i^{\text{cost+BTA}} = L_i^{\text{demand}} = L_i^{\text{supply}} < 0$.

Proof. Following the proof of Proposition 1, country k's first-order condition from (2) now is $p(X) - C_k'(x_k) - \tau_k\theta_k = 0$, where $\tau_k = \tau_i(\tau, \lambda_i)$ for $k = i, j$. Write the equilibrium outputs in terms of i's unilateral policy as $(x_i(\lambda_i), x_j(\lambda_i))$. At equilibrium, k's first-order condition from (2) becomes:

$$p(x_i(\lambda_i) + x_j(\lambda_i)) - C_k'(x_k(\lambda_i)) - \tau_k\theta_k = 0.$$

Differentiating this condition for i yields:

$$p'(x_i(\lambda_i) + x_j(\lambda_i))\left(\frac{dx_i}{d\lambda_i} + \frac{dx_j}{d\lambda_i}\right) - C_i''(x_i(\lambda_i))\frac{dx_i}{d\lambda_i} - \theta_i\frac{d\tau_i}{d\lambda_i} = 0,$$

which can be rearranged as:

$$[p'(\cdot) - C_i''(\cdot)]\frac{dx_i}{d\lambda_i} + p'(\cdot)\frac{dx_j}{d\lambda_i} - \theta_i\frac{d\tau_i}{d\lambda_i} = 0$$

and then solved in terms of i's output response:

$$\frac{dx_i}{d\lambda_i} = -\frac{-p'(\cdot)}{[-p'(\cdot) + C_i''(\cdot)]}\frac{dx_j}{d\lambda_i} - \frac{\theta_i}{[-p'(\cdot) + C_i''(\cdot)]}\frac{d\tau_i}{d\lambda_i},$$

where a symmetric expression holds for j:

$$\frac{dx_j}{d\lambda_i} = -\frac{-p'(\cdot)}{\left[-p'(\cdot) + C_j''(\cdot)\right]}\frac{dx_i}{d\lambda_i} - \frac{\theta_j}{\left[-p'(\cdot) + C_j''(\cdot)\right]}\frac{d\tau_i}{d\lambda_i}.$$

Solving the two last expressions simultaneously yields i's equilibrium output response:

$$\frac{dx_i}{d\lambda_i} = \frac{-\frac{1}{[-p'(\cdot) + C_i''(\cdot)]}\left[\theta_i - \frac{-p'(\cdot)}{\left[-p'(\cdot) + C_j''(\cdot)\right]}\theta_j\right]\frac{d\tau_i}{d\lambda_i}}{\left(1 - \frac{-p'(\cdot)}{[-p'(\cdot) + C_i''(\cdot)]}\frac{-p'(\cdot)}{\left[-p'(\cdot) + C_j''(\cdot)\right]}\right)}.$$

We are interested in the case where the policy is successful in reducing i's domestic emissions, $dx_i/d\lambda_i < 0$; this corresponds to $\theta_i/\theta_j > -p'(\cdot)/\left[-p'(\cdot) + C_j''(\cdot)\right] \equiv \underline{\theta} \in (0, 1)$, i.e., i's firms cannot be too clean relative to j's. (Otherwise the policy induces j's firms to contract output so strongly that this raises i's overall profitability of production.) Again a symmetric expression holds for j:

$$\frac{dx_j}{d\lambda_i} = \frac{-\frac{1}{\left[-p'(\cdot) + C_j''(\cdot)\right]}\left[\theta_j - \frac{-p'(\cdot)}{[-p'(\cdot) + C_i''(\cdot)]}\theta_i\right]\frac{d\tau_i}{d\lambda_i}}{\left(1 - \frac{-p'(\cdot)}{[-p'(\cdot) + C_i''(\cdot)]}\frac{-p'(\cdot)}{\left[-p'(\cdot) + C_j''(\cdot)\right]}\right)}.$$

Observe that $dx_j/d\lambda_i < 0$ if and only if $\theta_i/\theta_j < [-p'(\cdot) + C_i''(\cdot)]\,/[-p'(\cdot)] \equiv \overline{\theta}$, where $\overline{\theta} > 1$. (The same logic as for i applies also to j.) Combining i's and j's output responses gives an expression for the rate of output leakage:

$$-\frac{dx_j/d\lambda_i}{dx_i/d\lambda_i} = -\frac{\left[-p'(\cdot)\left(\frac{\theta_j}{\theta_i}-1\right)+\frac{\theta_j}{\theta_i}C_i''(\cdot)\right]}{\left[-p'(\cdot)\left(1-\frac{\theta_j}{\theta_i}\right)+C_j''(\cdot)\right]}.$$

From (4) in the proof of Proposition 1, we know that $C_k''(x_k) = \eta_k^S[p(X)/X](1/\sigma_k)$ and so, using the definitions $\varepsilon^D \equiv -p(\cdot)/Xp'(\cdot)$ and $\eta_k^S = 1/\varepsilon_k^S$, also $C_k''(x_k)/[-p'(X)] = (\varepsilon^D/\varepsilon_k^S)(1/\sigma_k)$. Therefore we can also write:

$$-\frac{dx_j/d\lambda_i}{dx_i/d\lambda_i} = -\frac{\left[\left(\frac{\theta_j}{\theta_i}-1\right)+\frac{\theta_j}{\theta_i}\frac{\varepsilon^D}{\varepsilon_i^S}\frac{1}{(1-\sigma_j)}\right]}{\left[\left(1-\frac{\theta_j}{\theta_i}\right)+\frac{\varepsilon^D}{\varepsilon_j^S}\frac{1}{\sigma_j}\right]}.$$

The formula for carbon leakage $L_i = (\theta_j/\theta_i)[-(dx_j/d\lambda_i)/(dx_i/d\lambda_i)]$ follows immediately. Finally, it is easy to verify that if $\theta_i = \theta_j$, for which the condition $\theta_i/\theta_j \in (\underline{\theta}, \overline{\theta})$ is always met, the formula simplifies to that of Proposition 2, so that $L_i^{\text{cost+BTA}} = L_i^{\text{demand}} = L_i^{\text{supply}} < 0$ as claimed. ■

A cost-raising policy with a BTA has the *direct* effect of raising both i's and j's total marginal cost, according to their respective emissions intensities—which induces *lower* output and emissions from each country. In addition, there is the *indirect* effect that each country's output cut is mitigated by competitive gain arising from this reduction in the other country's output—which induces *higher* output and emissions from each country.

The additional assumption that i is not too much cleaner than j, with $\theta_i/\theta_j > \underline{\theta} \in (0,1)$, ensures that the direct effect outweighs for i—so its policy is successful at reducing i's domestic emissions.

The flipside is that the BTA turns carbon leakage negative as long as j is not too much cleaner than i, with $\theta_i/\theta_j < \overline{\theta}$ (where $\overline{\theta} > 1$).

For the special case where i's and j's firms have identical emissions intensities, leakage is identical to that of demand-reducing and supply-raising policies from Proposition 2, $L_i^{\text{cost+BTA}} = L_i^{\text{demand}} = L_i^{\text{supply}}$. Intuitively, firms both in both countries are then hit by an identical cost shock which is economically equivalent to them being hit by an identical demand or supply shock.

Online Appendix C: Further details on policy illustrations

In this appendix, we provide additional details on the various policies that we describe in Section 5.

Policies in the EU ETS area

Cost-raising policies

Electricity

As discussed in the main text, the Dutch government announced a national CPF for the electricity sector in 2018. It is slated to increase from EUR 12.30/tCO_2 in 2020 to EUR 31.90/tCO_2 in 2030. In 2013, Great Britain introduced a carbon fee for its power sector, but we are not aware of any intra-EU leakage estimates. Such policies, if binding, suffer from intra-EU leakage as domestic generation gets replaced with imports. Table 1 in Frontier Economics (2018) estimates that the Dutch price floor will reduce domestic emissions by 26 million tCO_2 in 2030, but the net EU-wide emissions reduction is only 4 million tCO_2, implying $L = 0.85$. Vollebergh (2018) estimates internal carbon leakage to be 85% for the Dutch price floor and 61% for a regional CPF including the Benelux, France and Germany. We expect internal carbon leakage to be lower for the Great Britain's carbon fee as import supply is more inelastic due to interconnection constraints.

A national carbon price floor (CPF) is a "direct" cost-raising policy, but performance standards and mandates to reduce carbon in the electricity sector also fall into this category as they indirectly increase the domestic price tag on carbon. Examples include the British and Dutch policies to close their remaining coal-fired power plants by 2025 and 2030, respectively, and a similar plan is underway in Germany.[4]

Aviation

Aviation taxes are levied by several European countries, such as Austria, Germany, Norway and Sweden. Others, such as Denmark, Ireland and the Netherlands, abolished them after initial implementation. Such taxes are prone to leakage. For example, when the Netherlands adopted an aviation tax in July 2008 at a rate of EUR 11.25 for short-haul flight and EUR 45 for long-haul flights, about 50% of the decline in passengers at Dutch

[4]Sources: https://www.euractiv.com/section/energy/news/uk-government-spells-out-plan-to-shut-down-coal-plants/, https://cleantechnica.com/2018/05/22/the-netherlands-announces-ban-on-coal-plans-close-of-2-power-plants-by-2024/, https://www.businessgreen.com/bg/news/3026202/german-coalition-deal-holds-out-hopes-of-coal-phase-out-date.

airports was offset by increased passenger volumes in Belgium and Germany (Gordijn and Kolkman, 2011). They estimate that the tax accounted for nearly two million fewer passengers from Amsterdam's Schiphol Airport during the period over which the tax was in effect, while an extra one million Dutch passengers flew from foreign airports. As a result, the Dutch government abolished the tax in July 2009 although it currently considers reintroducing a very modest ticket tax of EUR 7 on all flights (Dutch News, 2018).

There is some evidence that aviation taxes are most likely levied in countries where leakage is mitigated. Figure 2 in PricewaterhouseCoopers (2017) shows that countries with large populations are more likely to have an aviation tax (France, Germany, Italy and the United Kingdom) as well as countries where the population is far away from low-tax airports abroad (e.g., Norway and Sweden). Austria is an exception to this rule given the proximity of Vienna to Bratislava. Greece, Croatia and Latvia—countries that also have aviation taxes—are relatively small, though their geographies are such that leakage may be less severe than for the Netherlands.

Supply-increasing policies

Germany and Spain have adopted some of the world's most ambitious incentives for wind and solar energy, which include feed-in tariffs and market premium programs. Abrell et al. (2019) find associated output and emissions leakage rates that are negative as a result of reduced imports through depressed wholesale prices. In their Table 3, they report d(import quantity)/d(policy) and d(domestic quantity)/d(policy), from which we calculate output leakage as -78%, -77%, -7% and -21% for German wind, German solar, Spanish wind and Spanish solar, respectively. Similarly, we compute carbon leakage from their Table 5: -49%, -50%, -6% and -19%, respectively. Averaged over wind and solar, we use $L = -0.50$ for Germany and $L = -0.12$ for Spain in Figure 2. The differences between output and emissions leakage in Germany and Spain suggest that the marginal unit of output reduction in Germany is approximately 50% more carbon intensive than the marginal reduction for its trading partners; for Spain the emissions intensity of these marginal units are about equal. Abrell et al. (2019), Table 3, shows that the German power mix is indeed dirtier than Spain's.

Policies in North America

California-Québec carbon trading

The California-Québec carbon market had an auction price floor of \$14.53 in 2018 and is scheduled to have a hard price ceiling of \$61.25 by 2021 (Politico, 2018). When the auction price floor binds, the unsold allowances are first placed in a holding account, from which they are re-introduced after two consecutive sold-out auctions (subject to a 25% volume limit per auction). If unsold for 24 months, they are moved to the Allowance Price Containment Reserve (APCR). In the case of the proposed hard price ceiling, the APCR will be practically infinite, so moving allowances into the APCR is essentially the same as retiring them.

Other key features of California's cap-and-trade policy include the requirement that allowances must be surrendered for emissions embodied in imported electricity and provisions to prevent resource shuffling. Resource shuffling is defined as "any plan, scheme, or artifice to receive credit based on emissions reductions that have not occurred, involving the delivery of electricity to the California grid" (Caron et al., 2015). If out-of-state generators can reconfigure transmission so that low-carbon electricity is diverted to California and high-carbon electricity to other states, even the border tax adjustment will not be able to mitigate leakage.

In the main text, we consider a counterfactual Western Climate Initiative (WCI) in which states surrounding California join the carbon market. The WCI (`http://www.wci-inc.org/`) started in 2007 as an initiative by the governors of Arizona, California, New Mexico, Oregon and Washington with a goal to develop a regional multi-sector cap-and-trade market. Most states left during the economic downturn in the early 2010s, but the idea of regional carbon trading has recently resurfaced in discussions among states. If California then imposed an additional unilateral carbon top-up fee, there will be emissions leakage to neighbouring states. There are of course no direct leakage estimates for a counterfactual policy, but estimates for similar policy settings do exist. Fowlie (2009) estimates that a carbon tax or cap-and-trade system in California that exempts out-of-state producers achieves only 25-35% of the total emissions reductions achieved under complete regulation (a carbon tax in Arizona, Nevada, New Mexico, Oregon, Utah and Washington). That is, $L = 0.65$-0.75. Caron et al. (2015) estimate that emissions leakage from California's cap-and-trade program (with border tax adjustments) varies between 9% when resource shuffling is successfully banned, and 45% when shuffling remains possible. Figure 2 plots the hypothetical California carbon top-up fee using $L = 0.09$ estimated in Caron et al. (2015), as their policy with a border tax adjustment and a ban on resource shuffling mimics California's current market rules most closely.

Regional Greenhouse Gas Initiative

The Regional Greenhouse Gas Initiative (RGGI) caps CO2 emissions from the power sector in ten Northeastern U.S. states. The program has a price floor; in 2012, the states decided to retire unsold allowances immediately. The policy also has a price ceiling. Several RGGI states have floated the idea of unilateral carbon policies; most notably, New York has proposed an additional carbon fee, which would not only apply in-state but also to imported electricity from other RGGI states. Shawhan et al. (2018) estimate that this policy leads to negative output leakage of $L = -0.77$. They assume a complete within-RGGI waterbed effect. They find a negative leakage rate for the following reason. RGGI exempts power plants with a capacity below 25 megawatts, while the New York carbon fee applies to all generation. This causes substitution from "NY non-RGGI" to "NY RGGI" sources. Due to the border tax adjustment, NY RGGI sources can increase output, thereby putting upward pressure on RGGI allowance prices and hence decreasing generation in other RGGI states.

We further note that Fell and Maniloff (2018) estimate substantial positive *external* carbon leakage ($L = 0.51$) from RGGI to non-RGGI states. The Shawhan et al. (2018) and Fell and Maniloff (2018) studies underscore that external and internal leakage are distinct phenomena that can even have different signs.

We do not plot New York's carbon fee policy in Figure 2 as we are not aware of an empirical estimate of the fraction of the time that the system trades at the price floor or ceiling, and thus W is missing.

References

Abrell, Jan, Mirjam Kosch, and Sebastian Rausch, "Carbon Abatement with Renewables: Evaluating Wind and Solar Subsidies in Germany and Spain," *Journal of Public Economics*, 2019, *169*, 172–202.

Baylis, Kathy, Don Fullerton, and Daniel H. Karney, "Leakage, Welfare and Cost-Effectiveness of Carbon Policy," *American Economic Review: Papers & Proceedings*, 2013, *103* (3), 332–337.

Beck, Ulrik and Peter K. Kruse-Andersen, "Endogenizing the Cap in a Cap-and-Trade System: Assessing the Agreement on EU ETS Phase 4," 2018. Working Paper.

Böhringer, Christoph, "Two Decades of European Climate Policy: A Critical Appraisal," *Review of Environmental Economics and Policy*, 2014, *8* (1), 1–17.

Borenstein, Severin, James Bushnell, and Frank A. Wolak, "California's Cap-and-Trade Market Through 2030: A Preliminary Supply/Demand Analysis," *Energy Institute at Haas Working Paper 281*, 2017.

_ , _ , _ , and Matthew Zaragoza-Watkins, "Expecting the Unexpected: Emissions Uncertainty and Environmental Market Design," *Energy Institute at Haas Working Paper 274*, 2018.

Caron, Justin, Sebastian Rausch, and Niven Winchester, "Leakage from Sub-national Climate Policy: The Case of California's Cap-and-Trade Program," *The Energy Journal*, 2015, *36* (2), 167–190.

Dutch News, "Dutch Plan Air Ticket Tax of EUR7, Introduction Slated for 2021," December 7th 2018. Available at https://www.dutchnews.nl/news/2018/12/dutch-plan-air-ticket-tax-of-e7-introduction-slated-for-2021/.

Edenhofer, Ottmar, Christian Flachsland, and Lisa K. Schmid, "Decarbonization and EU ETS Reform: Introducing a Price Floor to Drive Low-Carbon Investments," in M. Angrick, C. Kühleis, J. Landgrebe, and J. Weiss, eds., *12 Years of European Emissions Trading in Germany. Stocktaking and Perspectives for Effective Climate Protection*, Metropolis: London, United Kingdom, 2018, pp. 207–232.

European Commission, "Publication of the Total Number of Allowances in Circulation in 2017 for the Purposes of the Market Stability Reserve under the EU Emissions Trading System Established by Directive 2003/87/EC," Technical Report C(2018) 2801 final, European Commission 2018.

European Parliament and Council, "Directive (EU) 2018/410 of the European Parliament and of the Council of 14 March 2018 Amending Directive 2003/87/EC to Enhance Cost-Effective Emission Reductions and Low-Carbon Investments, and Decision (EU) 2015/1814," *Official Journal of the European Union*, 2018, *19.3.2018*, L76/4 – L76/27.

Fankhauser, Samuel, Cameron Hepburn, and Jisung Park, "Combining Multiple Climate Policy Instruments: How Not to Do It," *Climate Change Economics*, 2010, *1* (3), 209–225.

Fell, Harrison and Peter Maniloff, "Leakage in Regional Environmental Policy: The Case of the Regional Greenhouse Gas Initiative," *Journal of Environmental Economics and Management*, 2018, *87*, 1–23.

Flachsland, Christian, Michael Pahle, Dallas Burtraw, Ottmar Edenhofer, Milan Elkerbout, Carolyn Fischer, Oliver Tietjen, and Lars Zetterberg, "Five Myths about an EU ETS Carbon Price Floor," 2018. CEPS Policy Insights No 2018/17; available at http://aei.pitt.edu/95835/1/EU_ETS_Carbon_Price_Floor_Myths_and_enlightenment_final.pdf.

Fowlie, Meredith L., "Incomplete Environmental Regulation, Imperfect Competition, and Emissions Leakage," *American Economic Journal: Economic Policy*, 2009, *1* (2), 72–112.

Frontier Economics, "Research on the Effects of a Carbon Price Floor," 2018. Available at https://www.frontier-economics.com/media/2240/reasearch-effects-carbon-price-floor.pdf.

Gordijn, Hugo and Joost Kolkman, "Effects of the Air Passenger Tax: Behavioral Responses of Passengers, Airlines and Airports," 2011. KiM Netherlands Institute for Transport Policy Analysis. Available at http://publicaties.minienm.nl/documenten/effects-of-the-air-passenger-tax-behavourial-responses-of-passen.

IPCC, "Fourth Assessment Report: Climate Change 2007 – Working Group III: Mitigation of Climate Change," 2007. Chapter 11.7: International Spillover Effects, Intergovernmental Panel on Climate Change: Geneva, Switzerland.

Klobasa, Marian and Frank Sensfuss, "CO2-Minderung im Stromsektor durch den Einsatz Erneuerbarer Energien in den Jahren 2012 und 2013," 2016. Umweltbundesamt. Available at `https://www.umweltbundesamt.de/publikationen/co2-minderung-im-stromsektor-durch-den-einsatz`.

Murray, Brian and Nicholas Rivers, "British Columbia's Revenue-Neutral Carbon Tax: A Review of the Latest "Grand Experiment" in Environmental Policy," *Energy Policy*, 2015, *86*, 674–683.

Newbery, David M., David M. Reiner, and Robert A. Ritz, "The Political Economy of a Carbon Price Floor for Power Generation," *The Energy Journal*, 2019, *40* (1).

Perino, Grischa, "New EU ETS Phase 4 Rules Temporarily Puncture Waterbed," *Nature Climate Change*, 2018, *8* (4), 262–264.

Pizer, William A, "Combining Price and Quantity Controls to Mitigate Global Climate Change," *Journal of Public Economics*, 2002, *85* (3), 409–434.

Politico, "California Pro Preview: Carbon Prices," November 16th 2018. Available at `https://www.politico.com/states/california/newsletters/politico-california-pro-preview/2018/11/16/carbon-prices-136174`.

PricewaterhouseCoopers, "The Economic Impact of Air Taxes in Europe: Germany," 2017. KiM Netherlands Institute for Transport Policy Analysis. Available at `https://a4e.eu/wp-content/uploads/2017/10/The-economic-impact-of-air-taxes-in-Europe-Germany-004.pdf`.

Quemin, Simon and Raphaël Trotignon, "Intertemporal Emissions Trading and Market Design: An Application to the EU-ETS," 2018. Centre for Climate Change Economics and Policy Working Paper No. 348; available at `http://www.lse.ac.uk/GranthamInstitute/wp-content/uploads/2019/01/working-paper-316-Quemin-Trotignon.pdf`.

Roberts, Marc J. and Michael Spence, "Effluent Charges and Licenses under Uncertainty," *Journal of Public Economics*, 1976, *5* (3-4), 193–208.

Shawhan, Daniel, Karen Palmer, and Paul Picciano, "Benefits and Costs of New York's Carbon Pricing Initiative," 2018. Resources for the Future. Available at `https://www.scribd.com/document/389362390/RFF-NYISO-Carbon-Price-Analysis`.

Vollebergh, Herman, "National Measures Complementary to EU ETS," 2018. PBL Netherlands Environmental Assessment Agency. Available at `https://www.eprg.group.cam.ac.uk/wp-content/uploads/2018/12/H.-Vollebergh_final.pdf`.

Yamazaki, Akio, "Jobs and Climate Policy: Evidence from British Columbia's Revenue-Neutral Carbon Tax," *Journal of Environmental Economics and Management*, 2017, *83*, 197–216.