FAERE
French Association
of Environmental and Resource Economists

# Working papers

## Linking Permit Markets Multilaterally

Baran Doda - Simon Quemin - Luca Taschini

WP 2019.17

Suggested citation:

B. Doda, S. Quemin, L. Taschini (2019). Linking Permit Markets Multilaterally.
*FAERE Working Paper, 2019.17.*

ISSN number: **2274-5556**

www.faere.fr

# Linking Permit Markets Multilaterally*

Baran Doda[1,2] Simon Quemin[1,2,3] Luca Taschini[1,2,4]

[1]Grantham Research Institute
[2]ESRC Centre for Climate Change Economics and Policy
*London School of Economics*

[3]Climate Economics Chair – *Paris-Dauphine University (PSL)*

[4]*University of Edinburgh Business School*

This version: August 2019

**Abstract**

We formally study the determinants, magnitude and distribution of efficiency gains generated in multilateral linkages between permit markets. We provide two novel decomposition results for these gains, characterize individual preferences over linking groups and show that our results are largely unaltered with strategic domestic emissions cap selection or when banking and borrowing are allowed. Using the Paris Agreement pledges and power sector emissions data of five countries which all use or considered using both emissions trading and linking, we quantify the efficiency gains. We find that the computed gains can be sizable and are split roughly equally between effort and risk sharing.

**Keywords:** Climate change policy; International emissions trading systems; Multilateral linking; Effort sharing; Risk sharing.

**JEL classification codes:** Q58; H23; F15.

---

*An earlier version (September 2017) was circulated under the title 'A Theory of Gains from Trade in Multilaterally Linked ETSs'. The authors' email addresses are L.B.Doda@lse.ac.uk, S.Quemin@lse.ac.uk and L.Taschini1@lse.ac.uk.

*«[E]conomic theory suggests [carbon] markets should merge, and over time they probably will merge because of benefits of consolidation, including stability and lower cost.*

*...*

*[A recent study on linking found] benefits accrued for all parties as the market grew, though some parties benefitted more than others ... Expanding the partnership can be welfare improving in total but it can have distributional effects ... [That is,] different countries may benefit more or less depending on the mix of buyers and sellers who join any given market.»*

**— Forbes, December 7, 2015**

# 1 Introduction

Markets for emission permits have become an important policy instrument in responding to the climate change externality. A patchwork of emissions trading systems (ETSs), covering almost a quarter of global emissions, are now operational in jurisdictions including the EU, Switzerland, South Korea, New Zealand, China as well as several US states and Canadian provinces. Many more are in the pipeline (ICAP, 2019; World Bank, 2019). For the most part, permits cannot be traded across systems and the observed autarky prices for permits, their variability and comovement differ significantly. If these ETSs can be integrated through linking, substantial cost savings can in principle become available due to increased efficiency and stability. The opening quotes recognize these savings and highlight that although linking can be beneficial to all partnering jurisdictions, thorny distributional issues may arise depending on who the linking partners are. Understanding the determinants of multilateral linkages and how benefits from linking are shared is therefore a natural question, yet research is so far limited (Mehling et al., 2018). In this paper we propose a general model to describe and rigorously analyze multilateral linking, which presents an opportunity for achieving the ambitious goals of the Paris Agreement cost-effectively (Bodansky et al., 2016).

Our analysis quantifies the efficiency gains from linking which accrue to an individual jurisdiction participating in an arbitrary linkage group. We establish two decompositions of these gains. The first one identifies two independent sources of efficiency gains, namely effort- and risk-sharing gains. The former results from the heterogeneity in both abatement technologies and ambition levels of the group members. The latter is attributable to risk pooling associated with the uncertainty affecting each group member's demand for permits, our focus in this paper. The second decomposition allows us to express individual and aggregate gains in any linkage group as simple functions of aggregate gains in all its internal bilateral linkages, thereby representing a linkage group as the union of its building blocks.

These decompositions are complementary. The decomposition into effort and risk sharing offers a compact and intuitive interpretation of individual efficiency gains as a function of the expectation and variance of the autarky-linking price difference. Yet, it is unclear prima facie how efficiency gains, especially those due to risk sharing, relate to jurisdictions' characteristics. In this respect, the decomposition into internal bilateral links enables us to easily compute the efficiency gains generated by arbitrary linkage groups, which constitutes a pivotal tool for a quantitative illustration of our model. Additionally, it allows us to tease out and formally analyze the determinants of linkage gains and preferences.

Specifically, to study the efficiency gains from linking ETSs multilaterally and under uncertainty we start from a standard framework featuring permit demand shocks à la Weitzman (1974) and Yohe (1978). Our benchmark model is set up in a static environment where domestic emissions caps are assumed exogenously given and fixed to isolate the efficiency gains from linkage. The benchmark model abstracts from endogenous selection of domestic caps and intertemporal permit trading. We formally analyze the implications of allowing them in two extensions to our benchmark model below and show that our results continue to hold.

Our bilateral decomposition result allows us to rank groups from the perspective of individual jurisdictions and characterize the aggregate gains from the union of disjoint groups analytically. In turn, we emphasize why the conditions for the global market to be the most preferred group universally are unlikely to be satisfied in practice and we show that jurisdictional preferences for smaller linkage groups cannot be aligned without politically unpalatable compensatory monetary transfers. Additionally, we clarify the relationship between autarky and linking permit prices. In line with one's intuition, we show that relative to autarky linkage reduces price volatility on average though not necessarily for each individual entity. We provide a precise characterization of this effect.

We illustrate the quantitative implications of our model by focusing on all possible linkages across ETSs covering the $CO_2$ emissions from the power sectors of five real-world jurisdictions which all use or have considered both emissions trading and linking. Specifically, we calibrate our model to Australia, Canada, the EU, South Korea and the USA assuming that each jurisdiction implements its Paris Agreement pledge. We find that the linkage group which includes all five jurisdictions can generate total efficiency gains of up to \$3.26 billion (constant 2005 US\$) per annum which are split approximately equally between effort sharing, \$1.58 billion, and risk sharing, \$1.68 billion. Despite generating the largest aggregate gains, we note that this group is not the most preferred option unanimously. In fact, it is not the most preferred option for any individual jurisdiction. For instance, the USA would gain the most in a linking group with Australia and Europe. This three-jurisdiction group would also lead to lower price variability than in the group where all jurisdictions are linked.

How are these results altered if jurisdictions anticipate the option of future linking when choosing their domestic emissions caps, or if unrestricted intertemporal permit trade is allowed? First, we endogenize domestic cap selection based on self interest and in anticipation of linking à la Helm (2003). We derive closed-form solutions for the induced strategic and damage welfare impacts from linkage. The signs and magnitudes of these impacts are ambiguous and depend on the modeling structure and parameter distributions, as the subsequent literature attests, e.g. Carbone et al. (2009) and Gersbach & Winkler (2011). Crucially, they exist independently of the efficiency gains we focus on here, justifying the omission of these effects from the benchmark model. Second, we show in a multi-period setting how the introduction of unrestricted intertemporal permit trading alters, but crucially does not eliminate, the efficiency gains due to linking. In our quantitative illustration we find that allowing for unrestricted intertemporal trading reduces the effort- and risk-sharing gains by about 30% and 60%, respectively.

Throughout we abstract from economic, political and other (in)direct costs of linking which could preclude linkages that are otherwise beneficial. For example, significant and persistent differences in jurisdictional ambition levels imply some jurisdictions, and some regulated firms within jurisdictions, are net permit buyers in mutually beneficial trades but which nonetheless trigger ongoing financial transfers. These financial transfers in the buying jurisdictions and firms, as well as the persistently stricter-than-cap emission and higher permit price levels in the selling jurisdictions, can face political resistance (Jaffe et al., 2009).

These and related considerations can be a motive for increasing jurisdictions' emission caps strategically, e.g. as a result of domestic political lobbying (Habla & Winkler, 2013; Marchiori et al., 2017). Moreover, properly accounting for the general equilibrium interactions between the price in a linked permit market and the rest of the economy (Babiker et al., 2004; Carbone et al., 2009; Böhringer et al., 2014b), as well as the interactions with pre-existing distortionary tax systems (Bovenberg & Goulder, 1996; Babiker et al., 2003; Barrage, 2019) can generate implicit costs due to relative price changes triggered by linking. In addition, there are also direct transaction costs associated with inter-jurisdictional trading which might mitigate efficiency gains from linking (Baudry et al., 2019), e.g. due to a hierarchical transaction network structure (Karpf et al., 2018) and a home bias (Hintermann & Ludwig, 2019).

In fact, the balance between the efficiency gains and linkage costs may be one reason why some jurisdictions are already linked (e.g. California and Québec) while other links are expected to take a long time to emerge (e.g. the EU and the Chinese national ETS). In this paper we exclusively study the efficiency gains from linkage not because we think the associated costs are negligible but because the efficiency gains provide a potent incentive for jurisdictions to try to overcome them.

First and foremost, our paper is related to the literature on the economics of linking which has primarily emphasized three sources of gains from linking agreements, namely price convergence, a cost-effective reallocation of abatement efforts and a reduction of price volatility (Stevens & Rose, 2002; Flachsland et al., 2009; Fankhauser & Hepburn, 2010; Pizer & Yates, 2015; Ranson & Stavins, 2016; Doda & Taschini, 2017; Rose et al., 2018; Quemin & de Perthuis, 2019). Our two decomposition results allow us to formalize and refine these arguments in a multilateral setup under uncertainty. Specifically, we offer a precise characterization of both effort-sharing and risk-sharing gains from linkage, qualifying the results in Newell & Stavins (2003) and Caillaud & Demange (2017), who respectively studied efficiency gains in using market-based instruments relative to command-and-control policies, and linking disjoint ETSs. Additionally, we utilize our bilateral decomposition result to get a better sense of linkage preferences and we further characterize permit price properties.

While our work is framed in the context of linking permit markets, it also relates to the use of efficiency-improving trading ratios within permit markets (Holland & Yates, 2015). It is similar in spirit to the multinational production-location decision studied in de Meza & van der Ploeg (1987) and the choice of decentralization in permit markets analyzed by Yates (2002). Additionally, our results can have implications for interconnections between other types of supply-control programs with transferable licenses (e.g. production or fishery quotas) and international trade (e.g. cross-border electricity trading or energy unions). For instance, our paper formalizes some risk-sharing features attributable to permit transferability that were first highlighted in a more general context by Krishna & Tan (1999). In this respect, it also relates to several recent studies focusing on efficient risk sharing through international finance (Callen et al., 2015) or power interconnections (Antweiler, 2016).

Strategic aspects of inter-jurisdictional permit trading have been widely studied since the seminal contributions by Carraro & Siniscalco (1993), Barrett (1994) and Helm (2003). These papers reach somewhat pessimistic conclusions about the prospects of self-enforcing international permit trading schemes, for example, that linking markets creates incentives to relax domestic caps. More recently, Habla & Winkler (2018) find that allowing for strategic delegation of domestic cap selection leads to less stringent caps and higher aggregate emissions when permit markets are linked. We deliberately focus on effort- and risk-sharing gains to make a counterpoint and contribute to the literature that finds more positive results due to linking. For example, Helm & Pichler (2015) find that technology transfers among countries tend to reduce aggregate emissions when permits are internationally tradable. Additionally, Holtsmark & Midttømme (2019) show that linking permit markets can induce greater low-carbon investments and gradually lower emission caps, while Caparrós & Péreau (2017) and Heitzig & Kornek (2018) discuss models of gradual expansion of linked markets.

The remainder is organized as follows. Section 2 presents the model and discusses the theoretical results. Section 3.1 provides a qualitative illustration in a three-jurisdiction world. Section 3.2 contains a calibrated quantitative illustration. Section 4 introduces two extensions: endogenous cap selection and intertemporal trading. Section 5 concludes. All numbered tables and figures are provided at the end. There are two appendices dealing with the analytical derivations and proofs (A) and the description of our calibration methodology (B).

# 2 Model

## 2.1 Economic environment

To keep the model parsimonious and within the canonical framework, we consider a standard model of competitive markets for emission permits designed to regulate uniformly-mixed pollution in several jurisdictions in the manner of Weitzman (1974) and Yohe (1978), or more recently as in Hoel & Karp (2002), Newell & Pizer (2003, 2008) and Habla & Winkler (2018). In practice, the canonical framework used in these papers and others implies that we make three assumptions.

First, markets for permits are analyzed in isolation from the rest of the economy, i.e. interactions with other domestic and international markets through changes in relative prices are absent. Also absent are the general equilibrium interactions with pre-existing distortionary tax systems. As highlighted in the Introduction the effects emerging through these channels can be significant. Second, jurisdictions' benefits from emissions are expressed as quadratic functional forms which can be viewed as local approximations of general specifications and were shown to trace their real-world counterparts well (Klepper & Peterson, 2006; Böhringer et al., 2014a). Third, uncertainty is introduced in the form of additive shocks affecting jurisdictions' unregulated emission levels. Our benchmark model is static and takes jurisdictional emissions caps as fixed and independent of the decision to link. In Section 4, we show that our main results continue to hold in two extensions where we (1) endogenize domestic cap selection based on self interest in anticipation of linking à la Helm (2003) and (2) allow for intertemporal trading which interacts with linking in a dynamic multi-period setting.

**Jurisdictions** There are $n$ jurisdictions and $\mathcal{I} = \{1, \ldots, n\}$ denotes the set of jurisdictions. Aggregate benefits from emissions in jurisdiction $i \in \mathcal{I}$ are a function of the jurisdiction-wide emissions level $q_i \geq 0$ and of the random variable $\theta_i$ such that

$$B_i(q_i; \theta_i) = (\beta_i + \theta_i)q_i - q_i^2/(2\gamma_i), \tag{1}$$

where the parameters $\beta_i > 0$ and $\gamma_i > 0$ control the intercept and slope of $i$'s linear marginal benefit schedule, respectively.[1] Specifically, the parameter $\gamma_i$ reflects $i$'s abatement technology at the margin, hereafter technology for short. Thus, when comparing two jurisdictions $i$ and $j$, $\gamma_i > \gamma_j$ means that $i$ has access to a lower-cost abatement technology than $j$.

Jurisdiction $i$'s laissez-faire emissions maximize its benefits and are given by

$$\tilde{q}_i = \gamma_i(\beta_i + \theta_i). \tag{2}$$

The shock $\theta_i$ thus affects $i$'s laissez-faire emissions. For analytical convenience and without loss of generality, we assume that shocks are mean-zero with constant variance and that they may be correlated across jurisdictions. Specifically, for any pair $(i, j)$ we let

$$\mathbb{E}\{\theta_i\} = 0, \ \mathbb{V}\{\theta_i\} = \sigma_i^2, \text{ and } \mathrm{Cov}\{\theta_i; \theta_j\} = \rho_{ij}\sigma_i\sigma_j \text{ with } \sigma_i \geq 0 \text{ and } \rho_{ij} \in [-1; 1]. \tag{3}$$

For instance, $\theta_i > 0$ may reflect a favorable shock that increases $i$'s benefits from emissions, and therefore, the laissez-faire emissions relative to baseline emissions $\bar{q}_i = \mathbb{E}\{\tilde{q}_i\} = \gamma_i\beta_i$.

**Emissions caps** The emissions cap profile $(\omega_i)_{i\in\mathcal{I}}$ is exogenous and fixed. Having domestic caps independent of the decision to link anchors the aggregate level of emissions and rules out strategic spillovers. This allows us to (1) have well-defined autarky outcomes that serve as references throughout, (2) isolate the efficiency gains from linkage, and (3) compare these

[1] Jurisdiction $i$'s benefits correspond to the aggregate benefits accruing to all firms located within its boundaries. Indeed, covered firms are all united by a uniform price on emissions, which causes their marginal benefits to equalize. By horizontal summation, individual marginal benefit curves can thus be combined into one aggregate marginal benefit curve. Therefore, only the efficiency side of linking is covered here and the intra-jurisdictional distributional aspects are outside the scope of the paper.

gains across linkages and jurisdictions in a meaningful way. We later relax this assumption in Section 4.1 and discuss its implications.

For clarity, we express caps as proportional to technology by an ambition parameter such that

$$\omega_i = \alpha_i \gamma_i, \text{ where } \alpha_i \in (0; \beta_i) \text{ for all } i \in \mathcal{I}, \tag{4}$$

which implies that jurisdictional caps are all – but not equally – stringent relative to baseline. In particular, notice the negative relationship between $\alpha_i$ and the level of ambition implicitly embedded in $i$'s domestic cap, i.e. $\omega_i \to \bar{q}_i$ when $\alpha_i \to \beta_i$.

**Autarky equilibria** Under autarky, jurisdictions comply with their own caps. We assume that $\theta_i > \alpha_i - \beta_i$ for all $i$ and shock realizations so as to focus on interior autarky equilibria exclusively. That is, there are weak restrictions on individual shocks such that domestic caps are always binding. Specifically, autarky permit prices are positive and read

$$p_i = \bar{p}_i + \theta_i > 0 \text{ for all } i \in \mathcal{I}, \tag{5}$$

where $\bar{p}_i = \beta_i - \alpha_i > 0$ denotes $i$'s expected autarky price and notice $\bar{p}_i$ is lower for jurisdictions with higher $\alpha_i$.[2] First, note that for a positive (resp. negative) shock realization $\theta_i$, $i$'s autarky price is above (resp. below) $\bar{p}_i$. Second, note that when autarky prices differ – whether it be due to differences in ambition measured by $\bar{p}_i$ or shock realizations – the aggregate abatement effort is not efficiently allocated among jurisdictions. In particular, cost-efficiency could be improved by shifting some abatement away from relatively high-ambition (resp. high-shock) to low-ambition (resp. low-shock) jurisdictions until autarky price differentials are eliminated. We now characterize and quantify how linkage performs such a function.

## 2.2 Multilateral linkage and market equilibrium

Let $\mathcal{G} \subseteq \mathcal{I}$ be a non-empty subset of $\mathcal{I}$. We call $\mathcal{G}$ a group and $\mathcal{G}$-linkage the linked permit market between all jurisdictions in group $\mathcal{G}$. An interior $\mathcal{G}$-linkage equilibrium consists of the $(|\mathcal{G}|+1)$-tuple $(p_{\mathcal{G}}, (q_{\mathcal{G},i})_{i \in \mathcal{G}})$, where $p_{\mathcal{G}}$ is the equilibrium permit price in the linked market and

[2] Lecuyer & Quirion (2013) and Goodkind & Coggins (2015) provide explicit treatments of corner solutions in related contexts and demonstrate they can be of importance. Appendix B describing the calibration for our quantitative illustration shows that the interior equilibria assumption is innocuous for our quantitative results since $\bar{p}_i > 2\sigma_i$ for all $i \in \mathcal{I}$ we study. That is, assuming the shocks are normally distributed, zero-price corners occur with less than 2.5% probability in autarky and, a fortiori, under linkage.

$q_{\mathcal{G},i}$ denotes jurisdiction $i$'s equilibrium level of emissions.[3] The equilibrium is characterized by the equalization of marginal benefits across jurisdictions in $\mathcal{G}$ and market clearing, that is

$$MB_i(q_{\mathcal{G},i};\theta_i) = \beta_i + \theta_i - q_{\mathcal{G},i}/\gamma_i = p_{\mathcal{G}} \text{ for all } i \text{ in } \mathcal{G}, \text{ and } \sum_{i\in\mathcal{G}} q_{\mathcal{G},i} = \Omega_{\mathcal{G}} = \sum_{i\in\mathcal{G}} \omega_i, \tag{6}$$

where $\Omega_{\mathcal{G}}$ denotes $\mathcal{G}$'s cap. Cost-efficiency requires that any jurisdiction abates in proportion to its own technology, i.e. $\tilde{q}_i - q_{\mathcal{G},i} = \gamma_i p_{\mathcal{G}}$. In particular, the $\mathcal{G}$-linkage equilibrium price can be expressed as the technology-weighted average of autarky prices, that is

$$p_{\mathcal{G}} = \bar{p}_{\mathcal{G}} + \hat{\Theta}_{\mathcal{G}}, \text{ with } \bar{p}_{\mathcal{G}} = \sum_{i\in\mathcal{G}} \gamma_i \bar{p}_i/\Gamma_{\mathcal{G}} \text{ and } \hat{\Theta}_{\mathcal{G}} = \sum_{i\in\mathcal{G}} \gamma_i \theta_i/\Gamma_{\mathcal{G}}, \tag{7}$$

where $\Gamma_{\mathcal{G}} = \sum_{i\in\mathcal{G}} \gamma_i$ measures $\mathcal{G}$'s technology. Additionally, jurisdictional net permit demands are proportional to technology and the difference between the autarky and prevailing linking prices, that is

$$q_{\mathcal{G},i} - \omega_i = \gamma_i(p_i - p_{\mathcal{G}}). \tag{8}$$

In particular, jurisdiction $i$ is a net permit importer (resp. exporter) under $\mathcal{G}$-linkage provided that $p_i > p_{\mathcal{G}}$ (resp. $p_i < p_{\mathcal{G}}$), i.e. the linking price is lower (resp. higher) than its autarky price. Ceteris paribus, this shows that $\mathcal{G}$-linkage is observationally equivalent to an increase (resp. decrease) in $i$'s effective cap relative to autarky.

## 2.3 Efficiency gains in multilateral linkages

Because aggregate emissions are invariant, the welfare impacts from linkage only stem from an improvement in cost-effectiveness, i.e. a reduction in the costs of meeting the group-wide emissions cap $\Omega_{\mathcal{G}}$.[4] Specifically, the economic efficiency gains accruing to $i$ under $\mathcal{G}$-linkage denoted $\delta_{\mathcal{G},i}$ correspond to the difference between $i$'s benefits under $\mathcal{G}$-linkage (inclusive of

[3]Specifically, we further assume all jurisdictional shocks are bounded from above such that zero-emissions corners do not occur as a result of a link. In Appendix B, we show that such corners are typically rare since $\beta_i > \bar{p}_{\mathcal{G}} + 2\mathbb{V}\{\hat{\Theta}_{\mathcal{G}}\}^{1/2}$ for all jurisdictions $i$ and groups $\mathcal{G}$ we study. Our focus on interior market equilibria is thus innocuous for our analysis and allows simplification in (1) computing expected gains from linkage as damages from aggregate emissions are constant and (2) determining the linking price uniquely.

[4]In Section 4.1, we characterize two additional welfare impacts when domestic caps are endogenous with the decision to link, namely a strategic effect and a damage effect. Assuming invariant caps shuts down these two components but is without loss of generality for our characterization of efficiency gains.

proceeds from permit trading in the linked market) and autarky, that is

$$\begin{aligned}\delta_{\mathcal{G},i} &= B_i(q_{\mathcal{G},i};\theta_i) - p_{\mathcal{G}}(q_{\mathcal{G},i}-\omega_i) - B_i(\omega_i;\theta_i)\\ &= (q_{\mathcal{G},i}-\omega_i)^2/(2\gamma_i) = \gamma_i(p_i-p_{\mathcal{G}})^2/2 \geq 0.\end{aligned} \tag{9}$$

It is a well-known result that with fixed caps, linkage is mutually beneficial, i.e. efficiency gains are always non-negative. We characterize these gains further in the following proposition.

**Proposition 1.** *Under $\mathcal{G}$-linkage, the expected efficiency gains accruing to jurisdiction $i \in \mathcal{G}$ can be decomposed into effort- and risk-sharing gains, namely*

$$\begin{aligned}\mathbb{E}\{\delta_{\mathcal{G},i}\} = \gamma_i\mathbb{E}\{(p_i-p_{\mathcal{G}})^2\}/2 &= \gamma_i\Big(\underbrace{\mathbb{E}\{p_i-p_{\mathcal{G}}\}^2}_{\text{effort sharing}} + \underbrace{\mathbb{V}\{p_i-p_{\mathcal{G}}\}}_{\text{risk sharing}}\Big)\Big/2\\ &= \gamma_i\Big((\bar{p}_i-\bar{p}_{\mathcal{G}})^2 + \mathbb{V}\{\theta_i-\hat{\Theta}_{\mathcal{G}}\}\Big)\Big/2.\end{aligned} \tag{10}$$

*Proof.* Relegated to Appendix A.1. □

Jurisdiction $i$'s expected efficiency gains from $\mathcal{G}$-linkage are proportional to the expectation of the square of the difference in autarky and $\mathcal{G}$-linkage prices, i.e. the square of the distance in autarky-linking prices.[5] Crucially, efficiency gains can be decomposed into two non-negative components.[6]

The effort-sharing component is proportional to the square of the *expected autarky-linking price wedge*, relates to the intra-group variation in domestic ambition levels (i.e. expected autarky prices) and is independent of the shock structure. Intuitively, the larger this wedge, the larger the gains associated with the equalization of jurisdictional marginal benefits *on average*. In practice, however, significant disparities in expected autarky prices can compromise the political feasibility of a link for two reasons. First, they imply sizeable, persistent and politically-unpalatable monetary transfers associated with permit flows across jurisdictions. Second, they may connote different preferences in terms of environmental ambition or role of the carbon price signal as a domestic climate policy instrument.

The risk-sharing component is proportional to the *variance of the autarky-linking price wedge*, relates to jurisdictional and $\mathcal{G}$-wide shock characteristics, and is independent of jurisdictions'

[5] This result is the analog of Proposition 4.1 in Caillaud & Demange (2017). Moreover, note that summing $\delta_{\mathcal{G},i} = (q_{\mathcal{G},i}-\omega_i)^2/(2\gamma_i)$ over $i \in \mathcal{G}$ and taking expectation yields the aggregate comparative advantage of decentralization w.r.t. centralization for uniformly-mixed pollutants in Yates (2002).

[6] Appendix A.1 clarifies the sources of the efficiency gains identified in Proposition 1 by explicitly characterizing the reduction in the costs of emissions control under linking.

ambition levels.[7] That is, provided realized shocks differ across partnering systems, linking induces a strictly positive gain compared to the case without uncertainty, which is a strict Pareto-improvement due to risk pooling. Intuitively, controlling for the intra-group variation in expected autarky prices, the larger the ex-post wedge in autarky and linking prices, the larger the gains due to risk sharing. For instance, all else equal, $i$ will prefer to be in linkage groups where the price happens to be high w.r.t. its expectation when $i$'s (counterfactual) domestic price would have been low w.r.t. its expectation, and vice versa.

Moreover, because the $\mathcal{G}$-linkage price is the technology-weighted average of autarky prices in members of $\mathcal{G}$, all else equal, it is primarily driven by jurisdictions with higher $\gamma$'s. Similarly, for jurisdictions of similar technology, it is largely determined by those jurisdictions whose permit demand is highly variable. Therefore, only considering the risk-sharing component of gains, one expects that high-$\gamma$ and high-$\sigma$ jurisdictions may prefer to link with *several* jurisdictions to augment their autarky-linking price distances. By contrast, low-$\sigma$ (resp. low-$\gamma$) jurisdictions may prefer to link exclusively with a *single* low-$\sigma$ (resp. high-$\gamma$) jurisdiction, for otherwise the influence of that jurisdiction on the link outcome is likely to be mitigated. We further discuss the complex dependence of linkage preferences on the correlation coefficients in the next section and illustrate it using a qualitative example in Section 3.1.

## 2.4 Bilateral decomposition of gains in multilateral linkages

Equation (10) offers a compact and intuitive interpretation of jurisdictional gains in terms of autarky-linking price distance. This clarifies the behavior of the effort-sharing component, but it remains unclear prima facie how the risk-sharing component relates to jurisdictional characteristics. To illuminate this further, we unpack Equation (10) and to focus momentarily on the determinants of the risk-sharing component, we assume identical ambition across jurisdictions so that autarky-linking price wedges arise only due to shocks, i.e. $p_i - p_{\mathcal{G}} = \theta_i - \hat{\Theta}_{\mathcal{G}}$. Substituting this into Equation (9) and using the definition of $\hat{\Theta}_{\mathcal{G}}$, we obtain

$$\delta_{\mathcal{G},i} = \gamma_i \Big( \sum_{j\in\mathcal{G}} \gamma_j (\theta_i - \theta_j) \Big)^2 \Big/ (2\Gamma_{\mathcal{G}}^2). \tag{11}$$

[7] In other words, risk-sharing gains are invariant, irrespective of how caps are selected. See Section 4.1.

Expanding the above and taking expectations then yields

$$\mathbb{E}\{\delta_{\mathcal{G},i}\} = \gamma_i\Big(\sum_{j\in\mathcal{G}}\gamma_j^2(\sigma_i^2+\sigma_j^2-2\rho_{ij}\sigma_i\sigma_j) \\ +\sum_{(j,k)\in\mathcal{G}^2}\gamma_j\gamma_k(\sigma_i^2+\rho_{jk}\sigma_j\sigma_k-\rho_{ik}\sigma_i\sigma_k-\rho_{ij}\sigma_i\sigma_j)\Big)\Big/(2\Gamma_{\mathcal{G}}^2). \tag{12}$$

For clarity of interpretation, we first consider the most elementary group $\mathcal{G}=\{i,j\}$, i.e. a bilateral linkage. Letting $\Delta_{\{i,j\}}=\delta_{\{i,j\},i}+\delta_{\{i,j\},j}$ denote the aggregate economic gains from $\{i,j\}$-linkage, Equation (12) simplifies and gives

$$\mathbb{E}\{\Delta_{\{i,j\}}\} = \gamma_i\gamma_j(\sigma_i^2+\sigma_j^2-\rho_{ij}\sigma_i\sigma_j)/(2\Gamma_{\{i,j\}}) \geq 0, \text{ and} \tag{13a}$$

$$\mathbb{E}\{\delta_{\{i,j\},i}\}/\mathbb{E}\{\delta_{\{i,j\},j}\} = \gamma_j/\gamma_i. \tag{13b}$$

Intuitively and as described further in Doda & Taschini (2017), the aggregate risk-sharing gains from $\{i,j\}$-linkage are (1) positive as long as jurisdictional shocks are imperfectly correlated and jurisdictional volatility levels differ, for otherwise the two jurisdictions are identical in terms of shock characteristics, (2) increasing in both jurisdictional volatilities and technology parameters, (3) higher the more weakly correlated jurisdictional shocks are, and (4) for a given group's technology, maximal when jurisdictions have identical technology. Additionally, note that aggregate gains are apportioned between jurisdictions in inverse proportion to technology parameters. This is so because, for a given volume of trade, the distance between the autarky and linking prices is greater in the higher-cost technology jurisdiction.

Returning to the general case of any $\mathcal{G}$-linkage, we could pursue a similar approach to compute $\mathbb{E}\{\delta_{\mathcal{G},i}\}$ as $i$'s expected gains from a bilateral linkage between $i$ and $\mathcal{G}\backslash\{i\}$. However, the nature of the entity $\mathcal{G}\backslash\{i\}$ becomes exceedingly complex as the cardinality of $\mathcal{G}$ increases.

In this respect, one of our contributions is to recognize that bilateral linkages constitute the building blocks of the multilateral linkage analysis. Specifically, in a given linkage group, we show that it is more convenient to express the associated quantities as a function of the group's internal bilateral linkage quantities. With the tacit convention that $\Delta_{\{i,i\}}=0$ for any $i$, we can state the following proposition.

**Proposition 2.** *Any $\mathcal{G}$-linkage can be decomposed into its internal bilateral linkages. That is, $\mathcal{G}$-linkage gains (inclusive of both effort- and risk-sharing components) accruing to jurisdiction $i \in \mathcal{G}$ write as function of the aggregate gains in all bilateral linkages within $\mathcal{G}$*

$$\delta_{\mathcal{G},i} = \sum_{j \in \mathcal{G}\setminus\{i\}} \left\{ \Gamma_{\mathcal{G}\setminus\{i\}} \Gamma_{\{i,j\}} \Delta_{\{i,j\}} - (\gamma_i/2) \sum_{k \in \mathcal{G}\setminus\{i\}} \Gamma_{\{j,k\}} \Delta_{\{j,k\}} \right\} \Big/ \Gamma_{\mathcal{G}}^2. \tag{14}$$

*The number of such internal bilateral links is triangular and equals $\binom{|\mathcal{G}|+1}{2}$.*

*Proof.* Relegated to Appendix A.2. □

Proposition 2 helps us tease out jurisdictional linkage preferences. Specifically, jurisdiction $i$ is better off linking with sets of jurisdictions such that on the one hand, the aggregate gains in bilateral links between $i$ and each jurisdiction in these sets are high, and on the other hand, the aggregate gains in bilateral links internal to these sets are low. Referring to the above description of the determinants of the risk-sharing gains in bilateral links, these desirable sets, from the perspective of $i$, should comprise of jurisdictions that are similar to each other, with higher $\sigma$ and $\gamma$ than $i$, and negatively correlated with $i$. At the extreme and considering only the risk-sharing component of gains, $i$ would ideally like to link with as many replicas of its most preferred bilateral linking partner as possible.

Additionally, summing Equation (14) over all $i \in \mathcal{G}$ gives

$$\Delta_{\mathcal{G}} = \sum_{i \in \mathcal{G}} \delta_{\mathcal{G},i} = \sum_{(i,j) \in \mathcal{G}^2} \Gamma_{\{i,j\}} \Delta_{\{i,j\}} / (2\Gamma_{\mathcal{G}}). \tag{15}$$

In words, the aggregate $\mathcal{G}$-linkage gains write as a technology-weighted sum of all gains from bilateral linkages within $\mathcal{G}$. This decomposition result permits a more practical formulation and quantification of gains generated by an arbitrarily large group. Moreover, it allows us to provide an intuitive description of the efficiency gains in linking disjoint groups of linked jurisdictions. Specifically, let $\mathcal{G}' \subset \mathcal{G}$ and $\mathcal{G}''$ be the complement of $\mathcal{G}'$ in $\mathcal{G}$, i.e. $\mathcal{G} = \mathcal{G}' \cup \mathcal{G}''$ and $\mathcal{G}' \cap \mathcal{G}'' = \varnothing$. Then, we can express the aggregate gains in $\mathcal{G}$ as a function of those in $\mathcal{G}'$ and $\mathcal{G}''$ by unpacking Equation (15), that is

$$\Delta_{\mathcal{G}} = \Big( \Gamma_{\mathcal{G}'} \Delta_{\mathcal{G}'} + \Gamma_{\mathcal{G}''} \Delta_{\mathcal{G}''} + \sum_{(i,j) \in \mathcal{G}' \times \mathcal{G}''} \Gamma_{\{i,j\}} \Delta_{\{i,j\}} \Big) \Big/ \Gamma_{\mathcal{G}}. \tag{16}$$

Note that the third term in the parenthesis captures the interaction among jurisdictions in $\mathcal{G}'$ and $\mathcal{G}''$, which is precisely the quantity we want to isolate. To do so, we denote the aggregate gains of merging groups $\mathcal{G}'$ and $\mathcal{G}''$ by $\Delta_{\{\mathcal{G}',\mathcal{G}''\}}$ and define them by $\Delta_{\{\mathcal{G}',\mathcal{G}''\}} = \Delta_{\mathcal{G}} - \Delta_{\mathcal{G}'} - \Delta_{\mathcal{G}''}$. With this definition, Appendix A.3 shows that

$$\mathbb{E}\{\Delta_{\{\mathcal{G}',\mathcal{G}''\}}\} = \Big( \sum_{(i,j)\in\mathcal{G}'\times\mathcal{G}''} \Gamma_{\{i,j\}}\mathbb{E}\{\Delta_{\{i,j\}}\} - \Gamma_{\mathcal{G}''}\mathbb{E}\{\Delta_{\mathcal{G}'}\} - \Gamma_{\mathcal{G}'}\mathbb{E}\{\Delta_{\mathcal{G}''}\}\Big)\Big/\Gamma_{\mathcal{G}} \geq 0, \tag{17}$$

which is non-negative given the mutually beneficial nature of linkage with fixed caps. That is, the aggregate expected gains from the union of disjoint groups is no less than the sum of the separate groups' aggregate expected gains.[8] This implies the standard result that $\mathcal{I}$-linkage – the global market – is the linkage arrangement that is conducive to the highest aggregate cost savings in complying with the aggregate cap $\Omega_{\mathcal{I}}$.

## 2.5 Risk-sharing and permit price properties under linkage

The $\mathcal{G}$-linkage price $p_{\mathcal{G}} = \bar{p}_{\mathcal{G}} + \hat{\Theta}_{\mathcal{G}}$ is composed of two terms. The former, $\bar{p}_{\mathcal{G}} = \mathbb{E}\{p_{\mathcal{G}}\} = \sum_{i\in\mathcal{G}} \gamma_i \bar{p}_i / \Gamma_{\mathcal{G}}$, is commensurate with the stringency of the group-wide cap relative to baseline emissions. It measures the marginal cost of abatement when the group-wide expected abatement effort is allocated cost-efficiently. The latter, $\hat{\Theta}_{\mathcal{G}} = \sum_{i\in\mathcal{G}} \gamma_i \theta_i / \Gamma_{\mathcal{G}}$, quantifies the price impact due to the variability of the stringency of the group's cap relative to laissez-faire emissions that would be consistent with a profile of realized shocks. Indeed, given $(\theta_i)_{i\in\mathcal{G}}$, the quantity $\sum_{i\in\mathcal{G}} \gamma_i \theta_i$ measures the difference in the group's laissez-faire and baseline emissions. Then, dividing it by the group-wide technology $\Gamma_{\mathcal{G}}$ gives the corresponding price impact.

Next, we characterize the features of linkage in terms of risk-sharing by analyzing the properties of the linking permit price variability. As defined formally in Appendix A.4, a partition $\mathcal{P}'$ of $\mathcal{I}$ is coarser than a partition $\mathcal{P}$ if $\mathcal{P}'$ can be obtained from $\mathcal{P}$ by some sequence of linkages between groups in $\mathcal{P}$.[9] With this terminology we can then state the following proposition.

[8] Specifically, linkage is superadditive given fixed caps. In a related context, Proposition 2 in Hennessy & Roosen (1999) shows that merging firms covered under a permit market is superadditive but an analytical description of the gains as in Equation (17) is not provided.

[9] By way of example, consider that $\mathcal{P} = \{\mathcal{G}_1, \ldots, \mathcal{G}_6\}$ with $\bigcup_{i=1}^{6} \mathcal{G}_i = \mathcal{I}$. Then $\mathcal{P}' = \{\mathcal{G}_1 \cup \mathcal{G}_2, \mathcal{G}_3, \ldots, \mathcal{G}_6\}$ and $\mathcal{P}'' = \{\mathcal{G}_1 \cup \mathcal{G}_2 \cup \mathcal{G}_3, \mathcal{G}_4, \mathcal{G}_5 \cup \mathcal{G}_6\}$ are both coarser than $\mathcal{P}$, and $\mathcal{P}''$ is also a coarsening of $\mathcal{P}'$.

**Proposition 3.** *Linkage reduces permit price volatility on average in groups and partitions, but not necessarily for each of their member jurisdictions. That is,*
*(a) Linkage diversifies risk since for any group* $\mathcal{G}$ *and partitions* $(\mathcal{P}, \mathcal{P}')$ *with* $\mathcal{P}'$ *coarser than* $\mathcal{P}$*, we have* $\Gamma_{\mathcal{G}}\mathbb{V}\{p_{\mathcal{G}}\}^{1/2} \leq \sum_{i\in\mathcal{G}} \gamma_i \mathbb{V}\{p_i\}^{1/2}$ *and* $\sum_{\mathcal{G}\in\mathcal{P}'} \Gamma_{\mathcal{G}}\mathbb{V}\{p_{\mathcal{G}}\}^{1/2} \leq \sum_{\mathcal{G}\in\mathcal{P}} \Gamma_{\mathcal{G}}\mathbb{V}\{p_{\mathcal{G}}\}^{1/2}$.
*(b) Under* $\mathcal{G}$*-linkage, only when shocks are independent does it hold that* $\text{p-lim}_{|\mathcal{G}|\to+\infty} p_{\mathcal{G}} = \bar{p}_{\mathcal{G}}$. *In particular, relative to autarky, linkage always reduces price volatility in higher volatility jurisdictions but may increase it in lower volatility jurisdictions.*

*Proof.* Relegated to Appendix A.4. □

Statement $(a)$ indicates that linkage improves shock absorption and reduces price volatility on average relative to autarky. In a given group, the linking price volatility is smaller than the technology-weighted average of autarky price volatilities. That is, the variability of the group's cap stringency is less than the one implied by its members' individual cap stringencies taken together. Importantly, this property extends to partitions: the coarser a partition, the more diversified the domestic shocks on average. Obviously, on the flip side, linking implies that relative to autarky jurisdictional emission levels are uncertain and contingent on own and linkage partners' shock realizations. This, however, can be desirable as it dampens the impact of domestic shocks much like a supply-side control in an hybrid instrument does by introducing some responsiveness in domestic caps.[10]

Although linkage-induced diversification guarantees that price volatility is reduced on average in a group, Statement $(b)$ indicates that (1) enlarging a group does not always imply lower price variability, which would be true only if domestic shocks were independent and (2) not every member jurisdiction necessarily experiences a reduction in price volatility w.r.t. autarky. On the one hand, relatively volatile jurisdictions always experience reduced price volatility w.r.t. autarky as domestic shocks are spread over a thicker market and thus better cushioned. On the other hand, because linkage also creates exposure to foreign shocks, relatively stable jurisdictions may face higher volatility relative to autarky. We emphasize that linkage is always preferred to autarky even when it leads to higher price variability domestically, i.e. despite the fact that a jurisdiction might 'import' additional volatility as a result of the link. This holds because jurisdictional regulators are risk-neutral and because domestic

[10]In the normative framework of instrument selection, much has been written about the problem of perfectly inelastic permit supply since the seminal contribution by Roberts & Spence (1976) and, in general, hybrid instruments have been shown to outperform both pure price and pure quantity instruments. In our setup linking increases the price elasticity of permit demand whereas a hybrid instrument increases the price elasticity of permit supply, typically by introducing steps in supply curves, see e.g. Fell et al. (2012a).

and foreign shocks are of a fundamental nature, i.e. they lead to permit price differences, the elimination of which is the very source of efficiency gains.

# 3 Illustrations

## 3.1 Qualitative illustration

In this section we illustrate our theoretical results in a stylized setup with three jurisdictions $i$, $j$ and $k$. Taking jurisdiction $i$'s perspective, we compare its linkage options graphically in Figure 1. Throughout we assume that jurisdictions have identical ambition levels, i.e. autarky permit prices are equal in expectation across jurisdictions, and effort-sharing gains are thus nil. The calibrated quantitative illustration in Section 3.2 relaxes this assumption and provides monetary evaluations of both effort- and risk-sharing gains.

We start by describing the key features of Figure 1. The axes are the same across the panels of the figure and measure $\gamma_j$ and $\gamma_k$ with respect to the innocuous normalization $\gamma_i = 1$. The dot in the center of each panel identifies the point where the jurisdictions' technologies are identical, i.e. $\gamma_i = \gamma_j = \gamma_k = 1$. Throughout we also refer to the case where $\sigma_i = \sigma_j = \sigma_k > 0$ and $\rho_{ij} = \rho_{ik} = \rho_{jk} = 0$ as the symmetric uncertainty benchmark (*SUB*).

In Panel 1a we rule out the possibility of all three jurisdictions linking together, i.e. $\{i, j, k\}$, to focus on the simpler case of $i$'s possible bilateral linkage groups, namely $\{i, j\}$ and $\{i, k\}$. In this case, the 45° line depicts the indifference frontier along which $\{i, j\}$ and $\{i, k\}$ generate the same risk-sharing gains for $i$.[11] Above the frontier $i$ prefers to link with $k$ because $k$ has a lower-cost abatement technology than $j$ does. All else constant, deviations from *SUB* such as $\sigma_i = \sigma_j < \sigma_k$ or $\rho_{ij} = 0 > \rho_{ik}$ distort the indifference frontier to the dashed curve. These deviations imply that $k$ is $i$'s preferred partner in a larger region of the $\{\gamma_j, \gamma_k\}$-space.

In Panel 1b we revert back to *SUB* but now allow for the formation of $\{i, j, k\}$ *in addition to* the bilateral links just discussed. First, observe that at the point of identical technologies, $i$ prefers $\{i, j, k\}$ to the bilateral linkages. This is to be expected because with $j$ and $k$ ex ante identical, $\{i, j, k\}$ is twice as large as the bilateral groups $i$ could form and therefore offers more abatement opportunities ex post.[12] Now note that $i$'s indifference point between $\{i, j, k\}$ and bilateral linkages (denoted by a diamond) implies $\gamma_i < \gamma_j = \gamma_k$. Indeed, given the restrictions implicit in *SUB*, it must be that $j$ and $k$ can each offer sufficiently cheaper

[11] The analytical expressions for the indifference frontiers are available from the authors upon request.

[12] See Doda & Taschini (2017) for a discussion of the effects of market size on risk-sharing gains.

abatement opportunities to $i$ to render bilateral linkages at least as rewarding as $\{i, j, k\}$.

Second, deviations from *SUB* which do not break symmetry between $j$ and $k$ would move the point of indifference along the 45º line. For example, $\sigma_i < \sigma_j = \sigma_k$ would move the point of indifference northeast, thereby expanding the region in which $\{i, j, k\}$ is the preferred option symmetrically around the 45º line, and vice versa. Additionally, Panel 1c shows the implications of breaking the symmetry implicit in *SUB*. The case depicted in this panel distorts the indifference frontier in favor of the bilateral group $\{i, k\}$ which is consistent with deviations from *SUB* such that $\sigma_i = \sigma_j < \sigma_k$ or $\rho_{ik} < \rho_{ij} = \rho_{jk} = 0$.

Finally, it is informative to characterize $j$ and $k$'s linkage preferences in the same $\{\gamma_j, \gamma_k\}$-space. Panel 1d superimposes the linkage indifference frontiers for the three jurisdictions under *SUB*. The dark grey area at the center represents the zone where $\{i, j, k\}$ is simultaneously preferred by all three jurisdictions and should thus endogenously emerge. Here, $\{i, j, k\}$ is every jurisdiction's best option only when the technology parameters $\gamma_j$ and $\gamma_k$ do not deviate much from $\gamma_i$. Under more general conditions than under SUB, it is not clear prima facie whether jurisdictional characteristics are such that $\{i, j, k\}$ is the most preferred linkage option for all jurisdictions simultaneously.

The light grey areas at the top and in the southwest corners of Panel 1d represent the zones where $i$ and $k$ respectively prefer $\{i, k\}$-linkage the most. These zones do not overlap, so $\{i, k\}$-linkage cannot form endogenously without compensatory transfers. More generally, if jurisdictional linkage preferences are not aligned, in the sense that there is at least one jurisdiction whose most preferred option is not $\{i, j, k\}$, then we can show that a jurisdiction's most preferred group cannot simultaneously be the favourite of every other jurisdiction in that group.[13] Only inter-jurisdictional transfers could change this. In a world where transfers often face significant political-economy obstacles and thereby prove unwieldy, this non-alignment result can in part explain why linkage negotiations do not readily deliver large linkage groups.

## 3.2 Quantitative illustration

In this section we explore our model quantitatively by considering linkages between hypothetical ETSs regulating the $CO_2$ emissions from the power sector of five real-world jurisdictions with different levels of ambition and which are subject to different shocks. We assume annual compliance without permit banking and borrowing across compliance periods and compute

[13] See Appendix A.5 for a proof in the general case.

the annual effort- and risk-sharing gains implied by our model.[14]

**Data description, model calibration and caveats** We provide an overview of our calibration strategy, which is detailed in Appendix B. We focus on five jurisdictions with similar levels of development and which all use, or have considered, *both* emissions trading and linking: Australia (AUS), Canada (CAN), the European Union (EUR), South Korea (KOR) and the United States (USA). We obtained estimates of the annual baseline emissions and marginal abatement cost curves (MACCs) for the power sector of these jurisdictions in 2030 from Enerdata, a private research and consulting firm whose clients include several national governments, UNDP and the European Commission. Based on the Ener-Blue scenario of the POLES model, the company also provided us with its estimates of the annual emission caps consistent with the achievement of the 2030 targets defined in the Nationally Determined Contributions as announced at the Conference of Parties in Paris.

Equipped with these caps and MACCs, we compute the expected autarky permit prices using our model which range from 27.1\$/t$CO_2$ in AUS to 113.7\$/t$CO_2$ in CAN. The annual baselines ($\bar{q}_i$), emission caps ($\omega_i$) and corresponding expected autarky permit prices ($\bar{p}_i$) are reported in Table 1, which also contains the linear intercepts ($\beta_i$) and technology coefficients ($\gamma_i$) we calibrate with a linear interpolation of MACCs in the vicinity of domestic caps.[15]

We calibrate the shock properties using the residuals from the regression of historical emissions on time and time squared with data from the International Energy Agency. These shocks capture the net effect of stochastic factors that may influence emissions and their associated benefits, e.g. business cycles, TFP shocks, jurisdiction-specific events, changes in prices of factors of production, weather fluctuations, etc. Table 2 provides the volatility of the autarky permit prices as measured by the coefficient of variation, as well as the pairwise shock correlations implied by our theory. We note that there is large cross-jurisdiction variation in autarky price variability and that there are instances where the correlation between shocks is negative (e.g. KOR and EUR) or effectively zero (e.g. KOR and CAN).

Before proceeding to a discussion of our results, we point out a number of theoretical and quantitative caveats. Theoretically, all strategic and general equilibrium considerations which

[14] Section 4.2 provides an extension to study the attenuation in both effort- and risk-sharing gains due to linking when unrestricted intertemporal trading is allowed for the empirically relevant parameter values.

[15] The parameter $\gamma_i$ compounds the productivity of $i$'s abatement technology and $i$'s volume of regulated emissions. As such, comparing the ratios $\gamma_i/\bar{q}_i$ can give us a sense of the ordering of the volume-adjusted costs of abatement opportunities at the margin in the vicinity of the domestic caps. For instance, Table 1 shows that AUS has the cheapest abatement opportunities whereas the most expensive ones are in EUR.

have a bearing on the costs of linking, and therefore are relevant for the equilibrium cap levels or optimal emissions, are left out by construction. In other words, our partial equilibrium approach does not account for the welfare impact of changes in relative prices within (say due to changes in power prices) and across (say due to changes in terms of trade) jurisdictions. Also, absent are potentially amplified/diminished distortions in labor and capital markets.

Quantitatively, the POLES model of Enerdata is in essence a black box for us but we take its results as given. Accordingly, our quantitative results are sensitive to the assumptions of that model and its Ener-Blue scenario, particularly regarding the productivity and availability of abatement technologies in 2030 which determine our jurisdictions' marginal benefit schedules. Finally, we limit the analysis to the power sector due to data and cost constraints. This is the most natural and relevant sector to consider as it consists of large stationary sources which are easy to regulate, thus always brought under ETS regulations first. Bringing in additional sectors would imply some degree of *intra*-jurisdictional effort and risk sharing which could reduce efficiency gains from *inter*-jurisdictional trading despite a larger market size overall. These caveats imply that the per-annum efficiency gains computed below should be taken as illustrative upper bounds and with a large grain of salt.

**Discussion** Thanks to the bilateral decomposition result in Proposition 2 we can adopt a combinatorial approach to quantifying the *annual monetary gains*, in constant 2005US$, accruing to every jurisdiction in every possible linkage group.

Proposition 1 shows the jurisdictional gains are composed of an effort-sharing component and a risk-sharing component. These are illustrated in Figure 2 using three possible partitions of the set of five jurisdictions. We denote them $5J$ (top panel) when all five jurisdictions are linked together, and $\overline{2J3J}$ (middle panel) and $\underline{2J3J}$ (bottom panel) which generate the *greatest* and *smallest* gains among the ten possible partitions where no individual jurisdiction is in autarky. $\overline{2J3J}$ and $\underline{2J3J}$ are given by $\{\{\text{AUS,USA}\}, \{\text{CAN,EUR,KOR}\}\}$ and $\{\{\text{AUS,EUR}\}, \{\text{CAN,KOR,USA}\}\}$. The figure shows (1) how a group's gains are shared among the member jurisdictions and (2) the sources of gains for each jurisdiction. The areas of the various rectangles are comparable throughout, and proportional to the magnitude of the gains.[16] The colors identify jurisdictions and shades of a color distinguish the effort-sharing (light) and risk-sharing (dark) gains. The dotted areas in the panels for $\overline{2J3J}$ and $\underline{2J3J}$ illustrate the foregone gains with respect to $5J$.

[16]The small squares are an exception, e.g. KOR's effort-sharing gains in $5J$, and indicate gains too small to be visible in the graph.

In $5J$ the aggregate effort-sharing gains amount to \$1.58 billion, and those associated with risk sharing are \$1.68 billion, totalling \$3.26 billion. Risk sharing is the dominant source of gains in all jurisdictions but AUS. At \$1.38 billion AUS's effort-sharing gains account for almost 90% of aggregate effort-sharing gains. This is not surprising because the expected autarky-linking price wedge in AUS is the largest (\$27.1 vs \$86.5 per $tCO_2$). Conversely, EUR captures the largest risk-sharing gains which amount to \$0.62 billion or just over a third of the aggregate risk-sharing gains.

The sum of the efficiency gains generated in {AUS,USA} and {CAN,EUR,KOR} of $\overline{2J3J}$ are smaller than $5J$'s by about 14%, or \$0.45 billion. Most of the loss is due to the decline in the risk-sharing gains. The expected linking prices are given by \$84.1 and \$93.0, respectively. The largest effort-sharing gains among all possible bilateral links occur in {AUS,USA} which is about 90% of effort-sharing gains in $5J$. This is true despite the fact that the expected autarky price difference between AUS (\$27.1) and CAN (\$113.7) is greater because assuming similar abatement technologies USA is a much larger market than CAN. Were AUS and CAN to link their markets, the expected linking price would be too low (\$45.4) relative to that in {AUS,USA} and result in significant effort-sharing gains going unrealized. Of course, effort-sharing gains are only part of the story and the figure shows the risk-sharing gains generated in {AUS,USA} are small. Similarly, the risk-sharing gains generated in {CAN,EUR,KOR}, although sizable, are smaller than those in $5J$. This is because the absence of AUS and USA from this group reduces risk diversification opportunities significantly.

The groups {AUS,EUR} and {CAN,KOR,USA} in $\underline{2J3J}$ generate low effort-sharing gains as evidenced by the large difference between the groups' expected linking prices, \$66.4 and \$93.5, respectively; and low risk-sharing gains by grouping positively correlated jurisdictions together and thereby forgoing the opportunity of leveraging the negative correlation between EUR and KOR. Together, these two linking groups imply that almost 40%, or \$1.25 billion, worth of efficiency gains go unrealized relative to $5J$.

After illustrating the group-wide gains and how they are distributed across jurisdictions and across effort- and risk-sharing in *three select partitions*, we now take an individual jurisdiction's perspective and discuss efficiency gains and permit price volatility it may expect in *all groups* where it is a member. The three panels of Figure 3 use USA as an example. The top and middle panels show the efficiency gains for USA and permit price volatility as a function of the number of members in the group.

First, observe that $5J$ is not the group that generates the largest gains for USA. In light of the previous section, we conclude that $5J$ will therefore not emerge naturally for these five jurisdictions, even though it would generate the largest gains in aggregate. Neither is it the case that $5J$ delivers the lowest price volatility for USA which obtains in the bilateral link with EUR. In fact, USA permit price volatility may increase relative to its autarky level (horizontal line in the middle panel). However, we emphasize that in our model an increase in permit price volatility relative to autarky does not have any negative implications, which for many jurisdictions in the real world can be an important consideration.

Second, there is not a monotonic relationship between the magnitude of efficiency gains and cardinality of a group. For example, adding EUR to {AUS,USA} increases USA's efficiency gains while adding KOR or CAN decreases them. Third, linkage preferences do not tally. While USA would gain the more from adding EUR to {AUS,USA}, AUS would rather have KOR or CAN join the bilateral group next as it would benefit AUS more.

Finally, the bottom panel illustrates the large variation in the two components of gains across groups including USA which are ordered so risk-sharing gains decline along the x-axis. In all groups where AUS is a member, USA enjoys significant effort-sharing gains driven by the large difference in expected autarky prices between AUS and the others. In groups with greater number of members, USA effort-sharing gains tend to be lower as they are more diluted across jurisdictions relative to {AUS,USA}. Risk-sharing gains also vary significantly across all groups. USA efficiency gains consist almost exclusively of risk-sharing gains in groups that do not include AUS (e.g. {EUR,KOR,USA}) and may be larger than effort-sharing gains in groups that do include it (e.g. {AUS,EUR,USA}). These observations underline the need for a model to evaluate the efficiency gains from linking ETSs multilaterally.

# 4 Extensions

## 4.1 Linking with endogenous cap selection

Our analysis of linkage in Section 2 assumes away strategic cap selection and takes domestic caps as given. This can be justified by reference to the domestic political-economy constraints that emerge from the complex internal negotiation processes which must render the resulting policies acceptable to a host of actors with divergent interests, see e.g. Flachsland et al. (2009). Deviating from one's cap is therefore costly. However, one may contend that the prospects of inter-jurisdictional permit trading will drive regulators to set their caps in anticipation of linking based on self interest.

In this case, Helm (2003) showed that jurisdictions which expect to be net sellers (resp. buyers) of permits on the linked market have an incentive to inflate (resp. reduce) their caps relative to autarky to maximize their gains from linking. This strategic aspect and attendant shift in aggregate emissions and damages imply additional welfare impacts which in turn could compromise the feasibility of linkage, and autarky may even welfare-dominate linkage. At a minimum, jurisdictional linkage preferences may be altered. Below we analyze how endogenizing cap selection affects our analysis of linkage.

In what follows, we make the conventional assumption that marginal damages are constant and let $\eta_i$ denote $i$'s marginal damage (Pizer, 2002; Newell & Pizer, 2003). This assumption is consistent with damages being determined by the global cumulative emissions since the beginning of the Industrial Revolution (Allen et al., 2009; Allen, 2016). Here, it implies that jurisdictional reaction functions are orthogonal. We also assume that domestic caps are selected non-cooperatively with Cournot-Nash conjectural variations.[17]

Under autarky jurisdiction $i$ sets its cap $\omega_{\mathcal{A},i}$ to maximize its benefits net of damages, which simply yields $\omega_{\mathcal{A},i} = \gamma_i(\beta_i - \eta_i)$, i.e. $\bar{p}_{\mathcal{A},i} = \eta_i$. This reflects the weak form of the international free-riding problem, i.e. the intercepts of the reaction functions imply higher emission levels than in the global optimum, as $i$ does not internalize the negative externality inflicted by its emissions upon others.[18] Socially-efficient caps satisfy the Lindahl-Samuelson condition, are lower than the Cournot-Nash ones and imply all jurisdictions face the same price $\bar{p}_{\mathcal{I}} = \sum_{i \in \mathcal{I}} \eta_i$ in expectation, which is congruent with a global social cost of carbon (Kotchen, 2018).

Under $\mathcal{G}$-linkage, endogenizing cap selection is congruent with a two-stage game where jurisdictions set their caps at stage one and inter-jurisdictional permit trading occurs at stage two, which is typically solved in subgame Nash perfection using backward induction (D'Aspremont et al., 1983). As shown in Appendix A.6 jurisdiction $i$'s cap in anticipation of $\mathcal{G}$-linkage becomes

$$\omega_{\mathcal{G},i} = (\Gamma_{\mathcal{G}} - \gamma_i)(\langle \eta \rangle_{\mathcal{G}} - \eta_i) + \omega_{\mathcal{A},i} \geq \omega_{\mathcal{A},i} \Leftrightarrow \eta_i \leq \langle \eta \rangle_{\mathcal{G}}, \quad (18)$$

where $\langle \eta \rangle_{\mathcal{G}} = \sum_{i \in \mathcal{G}} \eta_i / |\mathcal{G}|$ is the average marginal damage in $\mathcal{G}$. Under the prospects of forming a linkage group the weak form of the free-riding problem is magnified (resp. mitigated) for relatively low-damage (resp. high-damage) jurisdictions and in turn, inter-jurisdictional

[17] If caps are selected cooperatively within a group, the prospects of inter-jurisdictional trading are inconsequential for cap selection (Carbone et al., 2009). Our results would be qualitatively similar under alternative conjectural variations because marginal damages are constant (MacKenzie, 2011; Gelves & McGinty, 2016).

[18] Due to the linearity of damages our framework does not capture its strong form, i.e. the crowding-out of domestic abatement efforts (reaction functions are negatively sloped with quadratic damages) which will always be strategic substitutes in a pure emissions game.

permit trading has an ambiguous effect on aggregate pollution relative to autarky since

$$\sum_{i\in\mathcal{G}}[\omega_{\mathcal{G},i} - \omega_{\mathcal{A},i}] = \sum_{i\in\mathcal{G}} \gamma_i(\eta_i - \langle\eta\rangle_{\mathcal{G}}) \gtrless 0, \tag{19}$$

whose sign depends on the distributions of the $\eta_i$'s and $\gamma_i$'s.[19] In fact, there is no consensus in the literature on this matter. For instance, Holtsmark & Sommervoll (2012) and Lapan & Sikdar (2019) show that linkage increases aggregate emissions relative to autarky absent and present trade in other goods, respectively. Using a computable general equilibrium model, Carbone et al. (2009) show that the opposite situation is more likely to occur.

The equilibrium market price under $\mathcal{G}$-linkage with endogenous cap selection reads

$$p^*_{\mathcal{G}} = \langle\eta\rangle_{\mathcal{G}} + \hat{\Theta}_{\mathcal{G}}. \tag{20}$$

Note that the $\mathcal{G}$-linkage prices with fixed and endogenous caps in Equations (7) and (20) are identical up to a shift in their deterministic parts from $\sum_{i\in\mathcal{G}} \gamma_i\bar{p}_i/\Gamma_{\mathcal{G}}$ to $\sum_{i\in\mathcal{G}} \bar{p}_{\mathcal{A},i}/|\mathcal{G}|$ and that $i$ is a net seller in expectation i.f.f. $\eta_i \leq \langle\eta\rangle_{\mathcal{G}}$. Because endogenous cap selection does not alter price variability, it will a fortiori not affect risk-sharing gains from linkage. Specifically, Helm (2003) shows that with endogenous caps the welfare impacts from linkage can be decomposed into three components, namely the efficiency gains from inter-jurisdictional trading, the strategic effect as measured by the market value of the difference in cap choices under autarky and linking, and the damage effect of changes in aggregate emissions. In the following proposition, we offer a precise analytical characterization of these three components.

**Proposition 4.** *With endogenous cap selection, the expected welfare impacts from $\mathcal{G}$-linkage in jurisdiction $i$ can be decomposed into three components*

$$\begin{aligned}\mathbb{E}\{\delta^*_{i,\mathcal{G}}\} =& \gamma_i\Big(\overbrace{(\eta_i - \langle\eta\rangle_{\mathcal{G}})^2}^{\text{effort sharing}\geq 0} + \overbrace{\mathbb{V}\{\theta_i - \hat{\Theta}_{\mathcal{G}}\}}^{\text{risk sharing}\geq 0}\Big)\Big/2 \\ &+ \underbrace{\langle\eta\rangle_{\mathcal{G}}(\Gamma_{\mathcal{G}} - \gamma_i)(\langle\eta\rangle_{\mathcal{G}} - \eta_i)}_{\text{strategic effect}\gtrless 0} + \underbrace{\eta_i\sum_{j\in\mathcal{G}}\gamma_j(\langle\eta\rangle_{\mathcal{G}} - \eta_j)}_{\text{damage effect}\gtrless 0} \gtrless 0.\end{aligned} \tag{21}$$

*Proof.* Relegated to Appendix A.6. ☐

As with exogenous caps in Proposition 1, efficiency gains have effort-sharing and risk-sharing

[19] For instance, if $\eta_i = \eta\gamma_i$ for all $i$ then Equation (19) reduces to $\eta\sum_{(i,j)\in\mathcal{G}^2}(\gamma_i - \gamma_j)^2/(2|\mathcal{G}|) \geq 0$ but note that this result is reversed if we assume $\eta_i = \eta/\gamma_i$ for all $i$ instead.

subcomponents, which are both non-negative. Observe that the latter is independent of cap selection which justifies our focus on exogenous caps in Section 2. That said, the interplay between the three welfare components is intricate and the latter two effects can be positive or negative. The strategic effect is positive iff $\eta_i < \langle \eta \rangle_{\mathcal{G}}$ while the damage effect is proportional to the variation in aggregate emissions between autarky and linkage and thus hard to sign. Hence, with endogenous cap selection the net welfare effect of linkage is ambiguous and the literature is again not decisive on this matter. With partial equilibrium models, Godal & Holtsmark (2011) and Holtsmark & Sommervoll (2012) find that linkage is unlikely to yield welfare gains while Dijkstra et al. (2011) and Antoniou et al. (2014) find just the opposite. With general equilibrium models, Marschinski et al. (2012) and Böhringer et al. (2014a) show that linkage-induced effects on welfare are ambiguous in general.

The above establishes the validity of our central result, namely Proposition 1, under strategic selection of domestic caps in anticipation of linking. In passing, we note that Proposition 3 on the stochastic properties of permit prices is also unaltered, but that Proposition 2, which provides an alternative formulation of individual efficiency gains, no longer holds.

## 4.2 Linking with banking and borrowing

Most if not all emissions trading systems allow for some form of intertemporal trading, that is banking issued permits for future compliance or borrowing future permits for present compliance. In Section 2, we abstracted from banking and borrowing when characterizing efficiency gains due to linking. By providing emitters with the opportunity to rearrange emissions over time, intertemporal trading can in principle reduce the price variability under autarky which in turn should shrink the risk pooling potential left over to linkage. In this section, we quantify the size and determinants of the efficiency gains due to linking *with* unrestricted intertemporal trading. We find that banking and borrowing does not eliminate the efficiency gains due to linking, and in some cases may increase them. The result turns on the persistence of shocks over time, the discount factor and the planning horizon.

For simplicity, we consider a stylized model of unrestricted banking and borrowing, which abstracts from constraints on the amount of permits that can be banked or borrowed. Without loss of generality, we assume jurisdictions apply the same discount factor $\lambda$ and that their benefit functions are time invariant. Additionally, given our discussion in the previous section, we revert to exogenous caps and further assume they are constant over time. Allowing for intertemporal trading alters market equilibrium permit prices but crucially not the definition of per-period linkage gains in Equation (14). In a given group $\mathcal{G}$ and period $t$ we denote by $p^\star_{\mathcal{G},t}$ and $p_{\mathcal{G},t}$ the prices with and without intertemporal trading, respectively.[20] Substituting them into (14) then gives the corresponding linkage gains, which we respectively denote by $\delta^\star_{i,\mathcal{G},t}$ and $\delta_{i,\mathcal{G},t}$. For instance, $\delta^\star_{i,\mathcal{G},t} = \gamma_i(p^\star_{i,t} - p^\star_{\mathcal{G},t})^2/2 \geq 0$. That is, allowing for intertemporal trading alters, but does not neutralize, the efficiency gains from inter-jurisdictional trading. Below, we characterize $\delta^\star_{i,\mathcal{G},t}$ as well as the ordering of $\delta^\star_{i,\mathcal{G},t}$ and $\delta_{i,\mathcal{G},t}$.

Consider two adjacent time periods $t$ and $t+1$. We let $\theta_{i,t}$ and $\theta_{i,t+1}$ denote the corresponding shocks in jurisdiction $i$ and assume that unconditional expectations are normalized to zero, i.e. $\mathbb{E}\{\theta_{i,t}\} = \mathbb{E}\{\theta_{i,t+1}\} = 0$. To specify the expectation of $\theta_{i,t+1}$ conditional on $\theta_{i,t}$ we assume that the joint distribution of $(\theta_{i,t}, \theta_{i,t+1})$ follows a standard AR(1) process. That is, using the shorthand notation $\mathbb{E}_t\{\cdot\}$ to denote expectation conditional on all information available at period $t$, $\mathbb{E}_t\{\theta_{i,t+1}\} = \varphi_i\theta_{i,t}$ where $\varphi_i \in [-1; 1]$ denotes the shock persistence.

Under autarky, the permit price in jurisdiction $i$ in period $t$ without intertemporal trading is simply given by Equation (5), i.e. $p_{i,t} = \bar{p}_i + \theta_{i,t}$. With intertemporal trading, the standard no-arbitrage condition with discounting and uncertainty (Samuelson, 1971; Schennach, 2000) is satisfied

$$p^\star_{i,t} = MB_i(q_{i,t}; \theta_{i,t}) = \lambda\mathbb{E}_t\{MB_i(q_{i,t+1}; \theta_{i,t+1})\} = \lambda\mathbb{E}_t\{p^\star_{i,t+1}\}. \tag{22}$$

That is, the discounted permit price is a martingale. Additionally, invoking the tower rule, Equation (22) can be chained over time with any given horizon of length $h \in \mathbb{N}$ yielding

$$p^\star_{i,t} = MB_i(q_{i,t}; \theta_{i,t}) = \lambda\mathbb{E}_t\{MB_i(q_{i,t+1}; \theta_{i,t+1})\} = \cdots = \lambda^h\mathbb{E}_t\{MB_i(q_{i,t+h}; \theta_{i,t+h})\}. \tag{23}$$

We can solve for the period-$t$ equilibrium price and expected price path with intertemporal trading over the horizon $h$ using Equation (23) and overall market closure at $t+h$, yielding

$$\mathbb{E}_t\{p^\star_{i,t+z}\} = \lambda^{-z}p^\star_{i,t} \text{ for any } z \in [\![0; h]\!] \text{ with } p^\star_{i,t} = \Big((h+1)\bar{p}_i + \theta_{i,t}\Phi_i\Big)\Big/\Lambda, \tag{24}$$

where $\Phi_i = \sum_{z=0}^h \varphi_i^z$ and $\Lambda = \sum_{z=0}^h \lambda^{-z}$, and which reduces to $p^\star_{i,t} = p_{i,t}$ only when $\varphi_i = 1$ and $\lambda = 1$.[21] When $\lambda < 1$, the deterministic part of $p^\star_{i,t}$ is smaller than that of $p_{i,t}$ due to temporal effort sharing. In practice, some abatement is postponed because $(h+1)/\Lambda$ decreases with

[20] Period length is deliberately left unspecified and need only coincide with any multiple of a given compliance period length, usually a year.

[21] We assume that the permit bank carried into $t$ from $t-1$, $b_{i,t-1}$, is zero for simplicity and without loss of generality. Accounting for past banking only introduces an offset in the certain component of $p^\star_{i,t}$ and thus leaves risk-sharing gains and the martingale property unchanged, see Equation (A.39) in Appendix A.7.

$h$ and $\lambda^{-1}$.[22] Not surprisingly, intertemporal trading reduces price variability because

$$\mathbb{V}\{p^{\star}_{i,t}\} = (\Phi_i/\Lambda)^2 \mathbb{V}\{p_{i,t}\} \leq \mathbb{V}\{p_{i,t}\}, \tag{25}$$

since $\Phi_i/\Lambda \leq 1$ with equality only when $\varphi_i = 1$ and $\lambda = 1$, or when $h = 0$. Given the properties of $\Phi_i/\Lambda$, (1) the greater the shock persistence, the less intertemporal trading can dampen the price impact of shocks;[23] (2) the longer the time horizon, the more jurisdictions have flexibility to spread out shocks over time, resulting in lower price variability today; and (3) the lower the discount factor, the less marked the price impact of shocks today as firms prefer to pass on more of the shocks to future periods. Only in the limit as $h \to \infty$ or $\lambda \to 0$ is price variability nil and the intertemporally tradable quantity instrument closely mimics the outcomes of a price instrument.[24] That is, unrestricted banking and borrowing alone *cannot* in general absorb all contemporaneous price variability.[25]

Similarly, the $\mathcal{G}$-linkage equilibrium price in period $t$ without intertemporal trading is given by Equation (7), i.e. $p_{\mathcal{G},t} = \bar{p}_{\mathcal{G}} + \sum_{i \in \mathcal{G}} \gamma_i \theta_{i,t}/\Gamma_{\mathcal{G}}$, whereas with intertemporal trading it reads

$$p^{\star}_{\mathcal{G},t} = \Big( (h+1)\bar{p}_{\mathcal{G}} + \sum_{i \in \mathcal{G}} \gamma_i \theta_{i,t} \Phi_i / \Gamma_{\mathcal{G}} \Big) \Big/ \Lambda, \tag{26}$$

and in expectation grows at the discount rate $\lambda^{-1} - 1$ over the horizon $h$. This implies that the static analysis of efficiency gains in Section 2 remains valid with intertemporal trading if shocks are rescaled by $\Phi_i/\Lambda$ to account for optimal, unlimited banking and borrowing. We can then state the following proposition.

[22] Similarly to linking, intertemporal trading generates effort and risk sharing gains. In a deterministic setting, temporal vs. spatial effort sharing gains have been analyzed, see e.g. Stevens & Rose (2002).

[23] This is not mathematically precise but conveys the core intuition. See Appendix A.7 for details.

[24] In a seminal paper comparing price and quantity instruments for stock pollutants, Newell & Pizer (2003) hint at this result (see their footnote 7) but do not develop it formally as their analysis abstracts from banking and borrowing. Extensions with intertemporal trading quantify further quantify this result, see equation 9 in Fell et al. (2012b) or Newell et al. (2005) and Pizer & Prest (2016) for a particular focus on policy updating.

[25] Moreover, comparing Equations (7) and (24) shows that, from the perspective of $i$ intertemporal trading is observationally equivalent to linking with $h$ uncorrelated replicas of $i$ whose individual shocks are given by $\{\theta_{i,t} \sum_{s=0}^{z} \varphi_i^s / \sum_{s=0}^{z} \lambda^{-s}\}_{z=1,\ldots,h}$. In other words, we quantify how time periods and jurisdictions are observationally equivalent 'divisions' in pollution permit markets as first analyzed in Yates (2002).

**Proposition 5.** *With unrestricted intertemporal trading over a finite time horizon of length $h$, the efficiency gains due to $\mathcal{G}$-linkage accruing to jurisdiction $i$ in any period $t$ amount to*

$$\mathbb{E}\{\delta^\star_{i,\mathcal{G},t}\} = \gamma_i\Big(\underbrace{(h+1)^2(\bar{p}_i - \bar{p}_\mathcal{G})^2/\Lambda^2}_{\text{effort sharing}} + \underbrace{\mathbb{V}\{\theta^\star_{i,t} - \hat{\Theta}^\star_{\mathcal{G},t}\}}_{\text{risk sharing}}\Big)\Big/2 \gtrless \mathbb{E}\{\delta_{i,\mathcal{G},t}\}, \tag{27}$$

*where* $\hat{\Theta}^\star_{\mathcal{G},t} = \sum_{i\in\mathcal{G}} \gamma_i\theta^\star_{i,t}/\Gamma_\mathcal{G}$ *and* $\theta^\star_{i,t} = \theta_{i,t}\Phi_i/\Lambda$, *with* $\Phi_i = \sum_{z=0}^{h} \varphi_i^z$ *and* $\Lambda = \sum_{z=0}^{h} \lambda^{-z}$.

*Proof.* Relegated to Appendix A.7 □

Proposition 5 extends Proposition 1 to a dynamic setup where, in addition to linking, unrestricted intertemporal trading within horizon $h$ is allowed. Although effort-sharing gains due to linking always decline when intertemporal trading is allowed, risk-sharing gains can decrease or increase, resulting in non-negative efficiency gains which may be lower or higher than in the case with no intertemporal trading. As further discussed in Appendix A.7, the ordering of $\mathbb{E}\{\delta^\star_{i,\mathcal{G},t}\}$ and $\mathbb{E}\{\delta_{i,\mathcal{G},t}\}$ depends on the complex interaction between the time horizon $h$, the discount factor $\lambda$ and the shock properties $\{\sigma_i, \rho_{ij}, \varphi_i\}_{i,j\in\mathcal{G}}$.

In our quantitative example, we argue that the efficiency gains due to linking are attenuated but not eliminated when intertemporal permit trading is allowed. To that end, we first note that when $\lambda < 1$ and shocks are similarly persistent across jurisdictions, i.e. $\varphi_i \simeq \varphi < 1$ for all $i \in \mathcal{G}$, efficiency gains are always attenuated by intertemporal trading and the ratios of effort- and risk-sharing gains with and without intertemporal permit trading are given by

$$\left(\frac{\mathbb{E}\{p^\star_{i,t} - p^\star_{\mathcal{G},t}\}}{\mathbb{E}\{p_{i,t} - p_{\mathcal{G},t}\}}\right)^2 \simeq \left(\frac{(h+1)(1-\lambda^{-1})}{1-\lambda^{-h-1}}\right)^2 < 1, \tag{28a}$$

$$\text{and} \quad \frac{\mathbb{V}\{p^\star_{i,t} - p^\star_{\mathcal{G},t}\}}{\mathbb{V}\{p_{i,t} - p_{\mathcal{G},t}\}} \simeq \left(\frac{(1-\varphi^{h+1})(1-\lambda^{-1})}{(1-\varphi)(1-\lambda^{-h-1})}\right)^2 < 1. \tag{28b}$$

Table 2 reports the estimated shock persistence parameters when an AR(1) process is fitted to our data which are approximately equal to 0.8 for all jurisdictions, i.e. $\varphi \approx 0.8$, which is coincidentally identical to the value Newell & Pizer (2003) estimate and use. Moreover, power producers typically hedge production up to three years ahead, so $h = 3$ seems a reasonable first-pass value.[26] Finally, we take $\lambda = 0.9$ for the discount factor. Plugging in these values in Equation (28) we find that intertemporal trading eats away about 30% and 60% of the

[26] See e.g. Neuhoff et al. (2012) and Schopp et al. (2015) and references therein in the case of the EU ETS. This 3-year hedging is typically incomplete as producers keep opportunities open to exploit changes through time, which means efficiency gains are likely to be reduced by less than what Equation (28) measures.

effort- and risk-sharing gains presented in Section 3.2, respectively.[27] Notwithstanding the attenuation in efficiency gains, we note that under unrestricted banking and borrowing, Proposition 2 is unaltered and Proposition 3 holds up to the shock rescaling above.

Finally, we highlight some of the differences between the stylized theory of intertemporal trading just analyzed and how it operates in practice. First, our theory assumes unrestricted intertemporal trading. In reality, borrowing is almost never authorized and banking can be limited, either by regulation via holding limits or due to firm-level internal or managerial constraints. Fell et al. (2012b) show that these constraints matter: as soon as they are expected to bind, banking offers little flexibility in smoothing out shocks.[28] Second, observed price dynamics in the EU ETS and elsewhere suggest that banking strategies by firms are not optimal, which might inter alia be caused by regulatory uncertainty (Salant, 2016; Fuss et al., 2018). Third, firms may be rationally short-sighted for hedging purposes or because they are poorly informed about future supply and demand conditions (Neuhoff et al., 2012; Schopp et al., 2015; Quemin & Trotignon, 2019). Last but not least, some risk may be jurisdiction-specific and so not diversifiable using intertemporal trading. Although a more comprehensive treatment of these considerations is in order, they can be thought of as impinging on intertemporal trading opportunities, de facto leaving more scope for inter-jurisdictional trading.

# 5 Conclusion

This paper advances the frontier of research on permit market integration by proposing a general model to describe and analyze multilaterally-linked ETSs formally in a partial equilibrium setup. In our model, efficiency gains and permit prices in any linkage group are well-defined objects and we study their analytical properties. First, we identify the two independent components which constitute the efficiency gains in any multilateral linkage, namely the effort- and risk-sharing components. The former is determined by the inter-jurisdictional variation in domestic ambition levels and the latter is driven by the nature of the uncertainty affecting the demand for permits in individual jurisdictions. Second, we decompose any multilateral linkage into its internal bilateral linkages. That is, we characterize aggregate and individual gains in any linkage group as a weighted average of the aggregate gains in all bilateral links that can be formed among its constituents. This decomposition formula is a

[27] In practice, caps are declining over time, typically at a rate of 2% per annum. When this is the case, intertemporal trading implies that effort-sharing gains decrease by 32% relative to 28% with constant caps.

[28] This is qualitatively similar to a decrease in $h$ or Samuelson effect, see e.g. Parsons & Taschini (2013).

practical tool to compute the gains generated in arbitrary linkage groups. It further allows us to rank groups from the perspective of individual jurisdictions and characterize the aggregate gains from the union of disjoint groups analytically. Third, we clarify the relationship between autarky and linking prices and show that relative to autarky, linkage reduces price volatility on average but not necessarily for individual entities. Finally, we show that our key findings hold when domestic caps are selected strategically or when unrestricted intertemporal trading is allowed. In other words, risk-sharing gains from linkage are independent of cap selection and remain substantial even when banking and borrowing is permitted.

Linkages between ETSs have a key role to play in the successful, cost-effective implementation of the Paris Agreement. To shed light on the magnitude and distribution of efficiency gains from linkage, we use a quantitative application with jurisdictions of similar levels of development which use, or have considered, both emissions trading and linking. We calibrate our model to the power sector $CO_2$ emissions of Australia, Canada, the EU, South Korea and the USA under the assumption that each jurisdiction implements its Paris Agreement pledges. The aforementioned theoretical and quantitative caveats of our analysis notwithstanding, we find that linking the five jurisdictions together can generate aggregate effort- and risk-sharing gains which are sizable and approximately equal to each other in magnitude. This suggests our theoretical findings can have significant relevance in policy-oriented applications and help overcome the economic and political barriers to linking in practice.

# Acknowledgments

The authors gratefully acknowledge funding from the Grantham Foundation for the Protection of the Environment, the UK Economic and Social Research Council (ESRC) through the Centre for Climate Change Economics and Policy, and the Climate Economics Chair. Simon Quemin acknowledges funding from LABEX Louis Bachelier Finance and Sustainable Growth (ANR 11-LABX-0019 Project) under the 'Investments for the Future' programme, in accordance with Article 8 of the Assigment Agreements for Communication Assistance.

We received helpful comments from the editor and two anonymous reviewers; participants at the AEA, AERE, AFSE, CESifo, EAERE, EEA, FAERE, FSR and ISEFI annual conferences in 2017 and 2018; and seminar participants at MCC Berlin, LSE, the universities of Bologna, Cambridge, Edinburgh, Essen-Duisburg, Geneva, Madrid (IPP), Paris-Dauphine, Queen Mary London and Verona. While revising this paper we learned about the passing of Martin Weitzman who has been an inspiration to each of us over the years and who generously commented on an earlier version of this paper. We are forever in his debt. We would also like to thank Alejandro Caparrós, Carolyn Fischer, Lawrence Goulder, Daniel Heyen, Michael Pahle, John Parsons, Philippe Quirion, Robert Ritz, François Salanié, Alessandro Tavoni, Oliver Tietjen, Ralph Winkler and Andrew Yates for valuable insights and conversations at different stages of this project. Federica Buricco's help in creating Figures 1 and 2 was invaluable. Our quantitative application relies heavily on modeling results from Enerdata for which we are grateful. The usual disclaimers apply.

# Tables

Table 1: Annual baseline emissions ($\bar{q}_i$, $10^6 tCO_2$) and annual emissions caps ($\omega_i$, $10^6 tCO_2$) obtained from Enerdata. Calculated expected autarky permit prices ($\bar{p}_i$, 2005US\$/$tCO_2$), calibrated flexibility coefficients ($\gamma_i$, $10^3(tCO_2)^2$/2005US\$), linear intercepts ($\beta_i$, 2005US\$/$tCO_2$) and ambition coefficients ($\alpha_i = \omega_i/\gamma_i$, 2005US\$/$tCO_2$) obtained using Enerdata data.

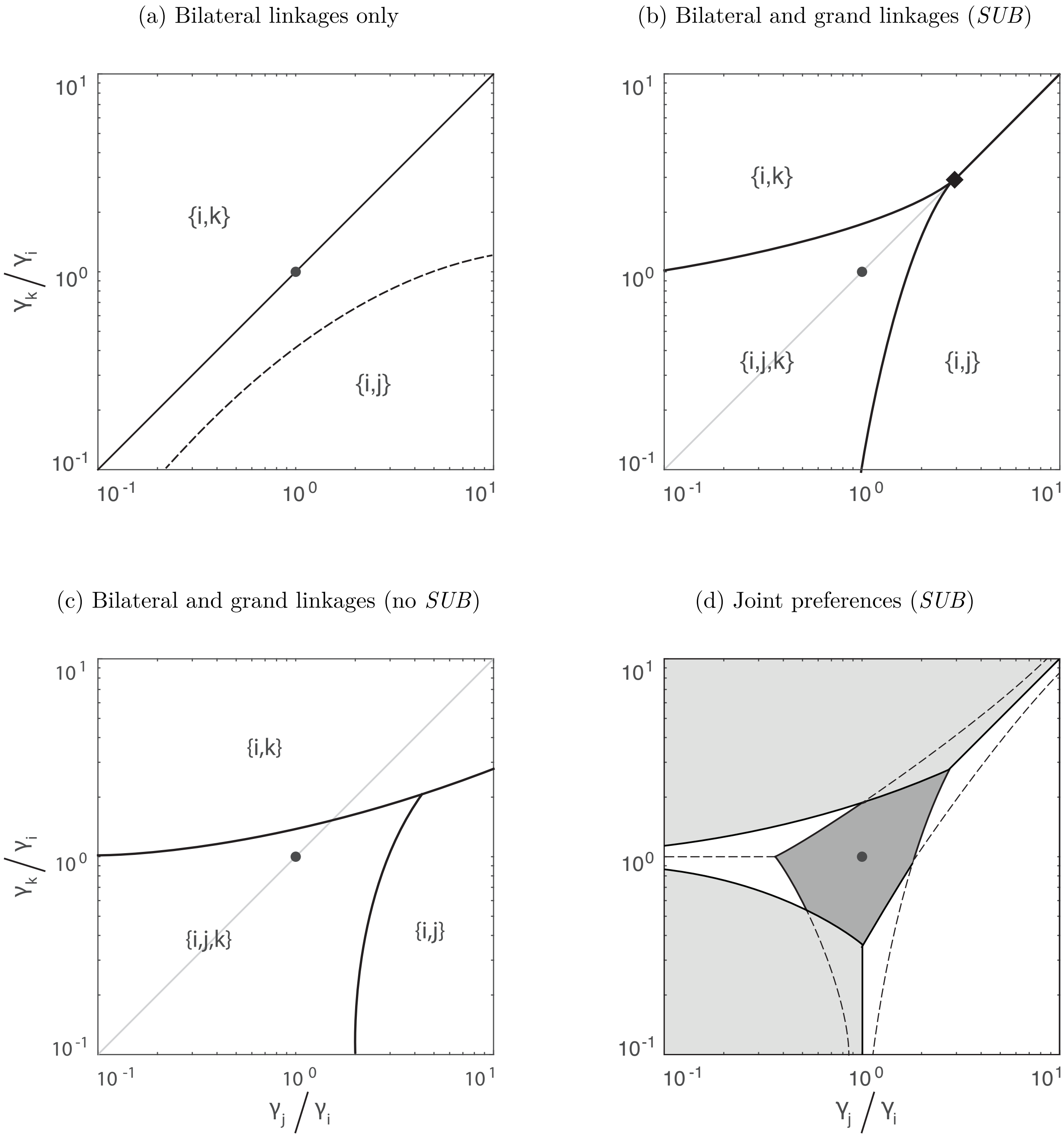
| | EUR | AUS | USA | CAN | KOR |
|---|---|---|---|---|---|
| $\bar{q}_i$ | 841.8 | 171.3 | 1,946.8 | 90.2 | 287.5 |
| $\omega_i$ | 724.1 | 150.1 | 1,469.3 | 66.3 | 225.8 |
| $\bar{p}_i$ | 89.8 | 27.1 | 92.8 | 113.7 | 92.6 |
| $\beta_i$ | 642.5 | 218.5 | 378.2 | 428.9 | 432.0 |
| $\gamma_i$ | 1,309.9 | 784.1 | 5,146.4 | 210.2 | 665.3 |
| $\alpha_i$ | 552.7 | 191.4 | 285.5 | 315.4 | 339.5 |
| $\gamma_i/\bar{q}_i$ | 1.6 | 4.6 | 2.6 | 2.3 | 2.3 |

Table 2: Coefficients of variation of autarky permit prices ($\sigma_i/\bar{p}_i$), pairwise correlation coefficients ($\rho_{ij}$) and AR(1) shock persistences ($\varphi_i$)

| | EUR | AUS | USA | CAN | KOR |
|---|---|---|---|---|---|
| $\sigma_i/\bar{p}_i$ | 0.35 | 0.44 | 0.24 | 0.49 | 0.52 |
| EUR | 1 | | | | |
| AUS | 0.36 | 1 | | | |
| USA | 0.07 | 0.42 | 1 | | |
| CAN | 0.18 | 0.18 | 0.43 | 1 | |
| KOR | -0.15 | 0.24 | 0.51 | 0.00 | 1 |
| $\varphi_i$ | 0.79 | 0.71 | 0.82 | 0.89 | 0.67 |

# Figures

Figure 1: Linkage preferences in the three-jurisdiction world $\{i, j, k\}$




(a) Bilateral linkages only

(b) Bilateral and grand linkages (*SUB*)

(c) Bilateral and grand linkages (no *SUB*)

(d) Joint preferences (*SUB*)

*Note:* *SUB* refers to the symmetric uncertainty benchmark defined in text.

Figure 2: Distribution and decomposition of efficiency gains in $5J$ (upper panel), $\overline{2J3J}$ (middle panel) and $\underline{2J3J}$ (lower panel). Colors identify jurisdictions and color shades identify risk- (dark) and effort-sharing (light) gains. The areas of the various rectangles are comparable across panels and proportional to the magnitude of the gains. Dotted areas demarcate foregone efficiency gains relative to $5J$. All numbers are billions of 2005US$ per year.

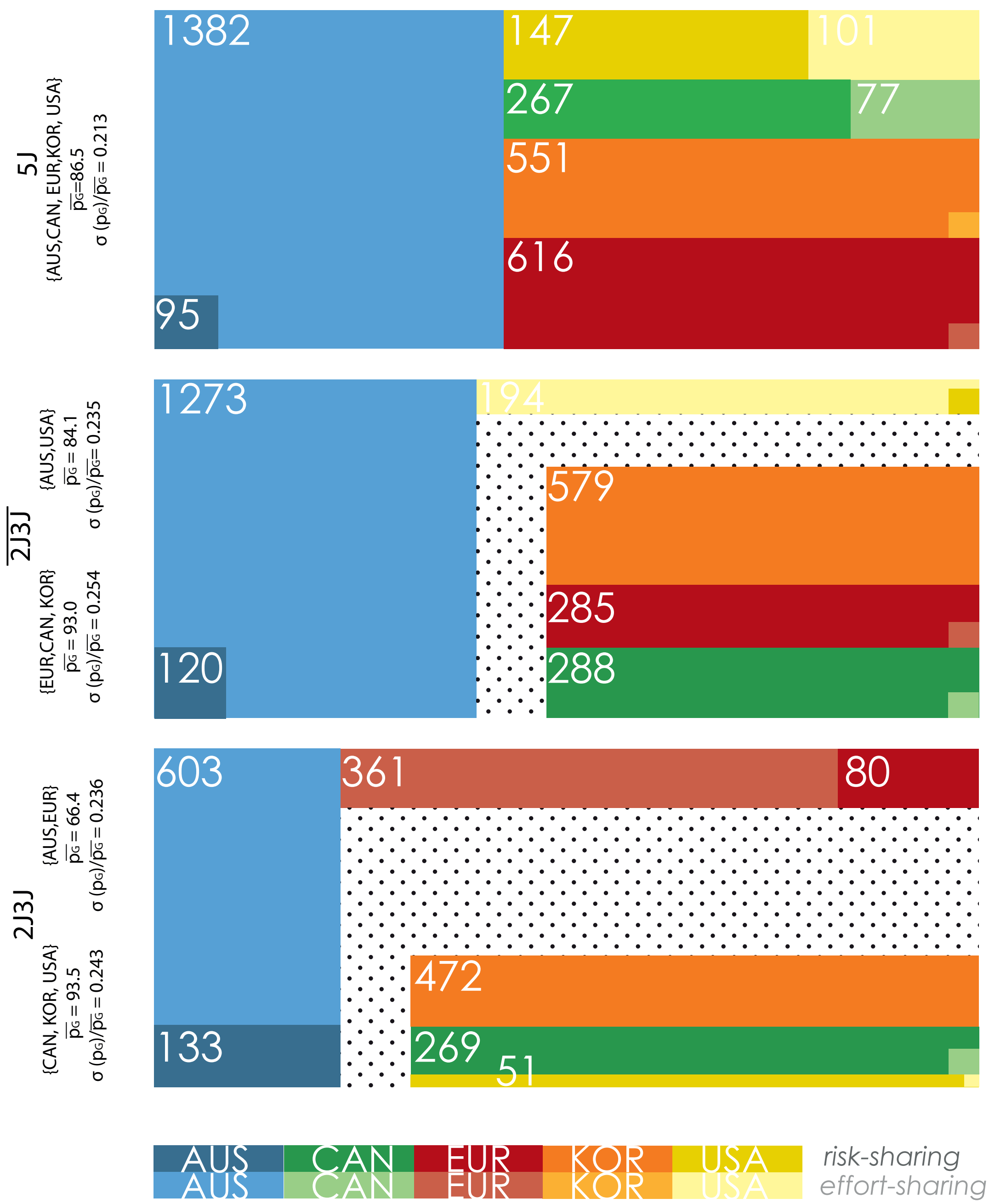
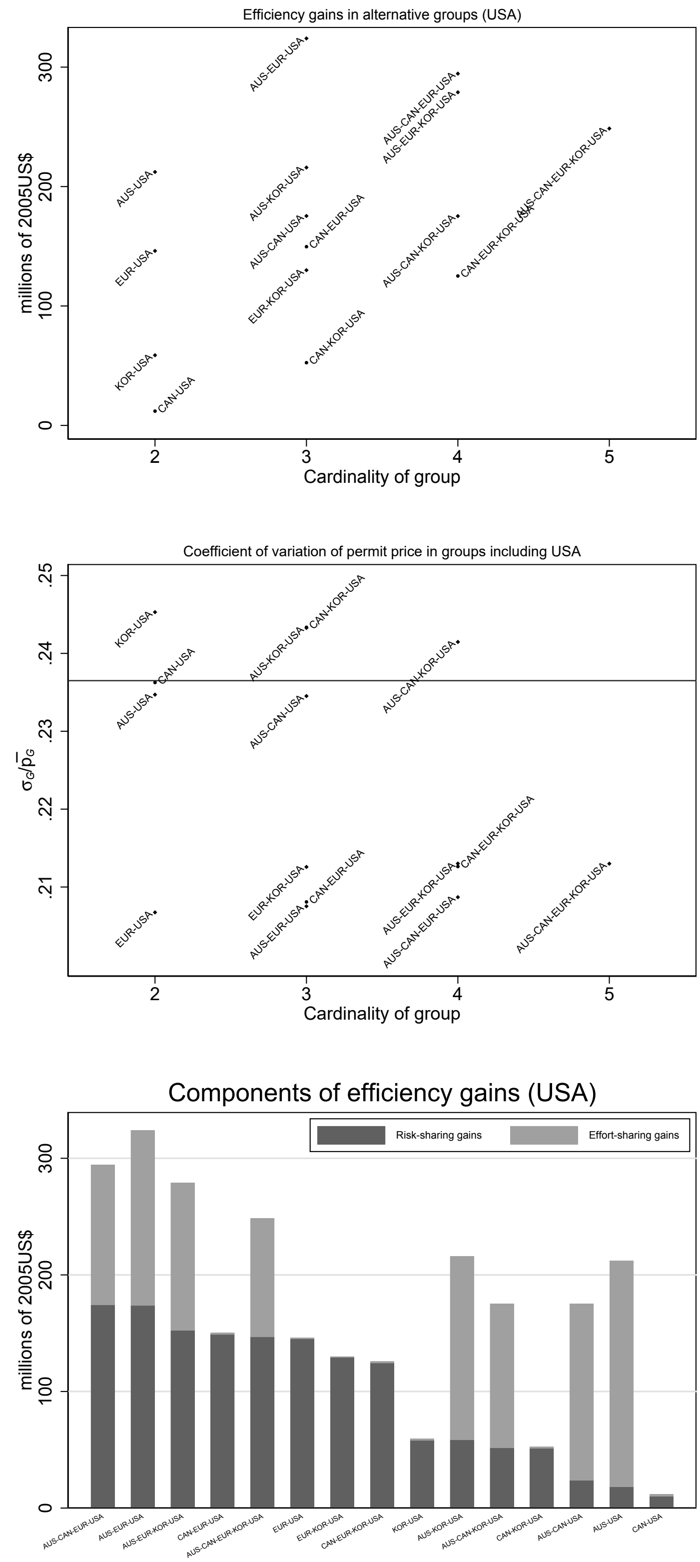


Figure 3: Expected per-annum efficiency gains, coefficients of variation of permit prices and components of gains in alternative linkage groups for USA




*Note:* The horizontal line in the middle panel indicates the coefficient of variation of USA autarky price.

# References


Allen, M. R. (2016). Drivers of Peak Warming in a Consumption-Maximizing World. *Nature Climate Change*, **6**, 684–6.

Allen, M. R., Frame, D. J., Huntingford, C., Jones, C. D., Lowe, J. A., Meinshausen, M. & Meinshausen, N. (2009). Warming Caused by Cumulative Carbon Emissions Towards the Trillionth Tonne. *Nature*, **458**, 1163–6.

Antoniou, F., Hatzipanayotou, P. & Koundouri, P. (2014). Tradable Permits vs Ecological Dumping when Governments Act Non-Cooperatively. *Oxford Economic Papers*, **66**(1), 188–208.

Antweiler, W. (2016). Cross-Border Trade in Electricity. *Journal of International Economics*, **101**, 42–51.

Babiker, M. H., Metcalf, G. E. & Reilly, J. (2003). Tax Distortions and Global Climate Policy. *Journal of Environmental Economics & Management*, **46**, 269–87.

Babiker, M. H., Reilly, J. & Viguier, L. (2004). Is International Emissions Trading Always Benficial? *Energy Journal*, **25**(2), 33–56.

Barrage, L. (2019). Optimal Dynamic Carbon Taxes in a Climate-Economy Model with Distortionary Fiscal Policy. *Review of Economic Studies*, in press.

Barrett, S. (1994). Self-Enforcing International Environmental Agreements. *Oxford Economic Papers*, **46**, 878–94.

Baudry, M., Faure, A. & Quemin, S. (2019). *Emissions Trading with Transaction Costs*. Working Paper, Grantham Research Institute, London School of Economics.

Baxter, M. & King, R. G. (1999). Measuring Business Cycles: Approximate Band-Pass Filters for Economic Time Series. *Review of Economics and Statistics*, **81**(4), 575–93.

Bodansky, D. M., Hoedl, S. A., Metcalf, G. E. & Stavins, R. N. (2016). Facilitating Linkage of Climate Policies Through the Paris Outcome. *Climate Policy*, **16**(8), 956–72.

Böhringer, C., Dijkstra, B. & Rosendahl, K. E. (2014a). Sectoral and Regional Expansion of Emissions Trading. *Resource & Energy Economics*, **37**, 201–25.

Böhringer, C., Lange, A. & Rutherford, T. F. (2014b). Optimal Emission Pricing in the Presence of International Spillovers: Decomposing Leakage and Terms-of-Trade Motives. *Journal of Public Economics*, **110**, 101–11.

Bovenberg, A. L. & Goulder, L. H. (1996). Optimal Environmental Taxation in the Presence of Other Taxes: General-Equilibrium Analyses. *American Economic Review*, **86**(4), 985–1000.

Burtraw, D., Palmer, K. L., Munnings, C., Weber, P. & Woerman, M. (2013). *Linking by Degrees: Incremental Alignment of Cap-and-Trade Markets*. Discussion Paper 13-04, Resources for the Future.

Caillaud, B. & Demange, G. (2017). Joint Design of Emission Tax and Trading Systems. *Annals of Economics & Statistics*, **127**, 163–201.

Calderón, C., Chong, A. & Stein, E. (2007). Trade Intensity and Business Cycle Synchronization: Are Developing Countries any Different? *Journal of International Economics*,

**71**(1), 2–21.

Callen, M., Imbs, J. & Mauro, P. (2015). Pooling Risk among Countries. *Journal of International Economics*, **96**(1), 88–99.

Caparrós, A. & Péreau, J.-C. (2017). Multilateral versus Sequential Negotiations over Climate Change. *Oxford Economic Papers*, 1–23.

Carbone, J. C., Helm, C. & Rutherford, T. F. (2009). The Case for International Emission Trade in the Absence of Cooperative Climate Policy. *Journal of Environmental Economics & Management*, **58**(3), 266–80.

Carraro, C. & Siniscalco, D. (1993). Strategies for the International Protection of the Environment. *Journal of Public Economics*, **52**, 309–28.

Christiano, L. J. & Fitzgerald, T. J. (2003). The Band Pass Filter. *International Economic Review*, **44**(2), 435–65.

D'Aspremont, C., Jacquemin, A., Gabszewicz, J. J. & Weymark, J. A. (1983). On the Stability of Collusive Price Leadership. *Canadian Journal of Economics*, **16**(1), 17–25.

Dijkstra, B. R., Manderson, E. & Lee, T.-Y. (2011). Extending the Sectoral Coverage of an International Emission Trading Scheme. *Environmental & Resource Economics*, **50**(2), 243–66.

Doda, B. (2014). Evidence on Business Cycles and $CO_2$ Emissions. *Journal of Macroeconomics*, **40**, 214–27.

Doda, B. & Taschini, L. (2017). Carbon Dating: When is it Beneficial to Link ETSs? *Journal of the Association of Environmental & Resource Economists*, **4**(3), 701–30.

Fankhauser, S. & Hepburn, C. (2010). Designing Carbon Markets, Part II: Carbon Markets in Space. *Energy Policy*, **38**(8), 4381–7.

Fell, H., Burtraw, D., Morgenstern, R. D. & Palmer, K. L. (2012a). Soft and Hard Price Collars in a Cap-and-Trade System: A Comparative Analysis. *Journal of Environmental Economics & Management*, **64**(2), 183–98.

Fell, H., MacKenzie, I. A. & Pizer, W. A. (2012b). Prices versus Quantities versus Bankable Quantities. *Resource & Energy Economics*, **34**(4), 607–23.

Flachsland, C., Marschinski, R. & Edenhofer, O. (2009). To Link or Not To Link: Benefits and Disadvantages of Linking Cap-and-Trade Systems. *Climate Policy*, **9**(4), 358–72.

Forbes (2015). How a Global Carbon Price is Slipping in the Back Door, December 7.

Fuss, S., Flachsland, C., Koch, N., Kornek, U., Knopf, B. & Edenhofer, O. (2018). A Framework for Assessing the Performance of Cap-and-Trade Systems: Insights from the European Union Emissions Trading System. *Review of Environmental Economics & Policy*, **12**(2), 220–41.

Gelves, A. & McGinty, M. (2016). International Environmental Agreements with Consistent Conjectures. *Journal of Environmental Economics & Management*, **78**, 67–84.

Gersbach, H. & Winkler, R. (2011). International Emission Permit Markets with Refunding. *European Economic Review*, **55**(6), 759–73.

Godal, O. & Holtsmark, B. (2011). Permit Trading: Merely an Efficiency-Neutral Redistribution away from Climate-Change Victims? *Scandinavian Journal of Economics*, **113**(4), 784–97.

Goodkind, A. L. & Coggins, J. S. (2015). The Weitzman Price Corner. *Journal of Environmental Economics & Management*, **73**, 1–12.

Habla, W. & Winkler, R. (2013). Political Influence on Non-Cooperative International Climate Policy. *Journal of Environmental Economics & Management*, **66**, 219–34.

Habla, W. & Winkler, R. (2018). Strategic Delegation and International Permit Markets: Why Linking May Fail. *Journal of Environmental Economics & Management*, **92**, 244–50.

Heitzig, J. & Kornek, U. (2018). Bottom-Up Linking of Carbon Markets: Farsightedness, Delayed Cap Coordination and Reversibility. *Nature Climate Change*, **8**, 205–9.

Helm, C. (2003). International Emissions Trading with Endogenous Allowance Choices. *Journal of Public Economics*, **87**(12), 2737–47.

Helm, C. & Pichler, S. (2015). Climate Policy with Technology Transfers and Permit Trading. *Environmental & Resource Economics*, **60**(1), 37–54.

Hennessy, D. A. & Roosen, J. (1999). Stochastic Pollution, Permits, and Merger Incentives. *Journal of Environmental Economics & Management*, **37**(3), 211–32.

Hintermann, B. & Ludwig, M. (2019). *Home Country Bias in International Emissions Trading: Evidence from the EU ETS*. Working Paper, University of Basel.

Hoel, M. & Karp, L. (2002). Taxes versus Quotas for a Stock Pollutant. *Resource & Energy Economics*, **24**(4), 367–84.

Holland, S. P. & Yates, A. J. (2015). Optimal Trading Ratios for Pollution Permit Markets. *Journal of Public Economics*, **125**, 16–27.

Holtsmark, B. & Sommervoll, D. E. (2012). International Emissions Trading: Good or Bad? *Economics Letters*, **117**(1), 362–4.

Holtsmark, K. & Midttømme, K. (2019). *The Dynamics of Linking Permit Markets*. Working Paper 7548, CESifo.

ICAP (2019). *Status Report 2019*. Berlin: International Carbon Action Partnership.

Jaffe, J., Ranson, M. & Stavins, R. N. (2009). Linking Tradable Permit Systems: A Key Element of Emerging International Climate Change Architecture. *Ecology Law Quarterly*, **36**(4), 789–809.

Karpf, A., Mandel, A. & Battiston, S. (2018). Price and Network Dynamics in the European Carbon Market. *Journal of Economic Behavior & Organization*, **153**, 103–22.

Klepper, G. & Peterson, S. (2006). Marginal Abatement Cost Curves in General Equilibrium: The Influence of World Energy Prices. *Resource & Energy Economics*, **28**(1), 1–23.

Kotchen, M. J. (2018). Which Social Cost of Carbon? A Theoretical Perspective. *Journal of the Association of Environmental & Resource Economists*, **5**(3), 673–94.

Krishna, K. & Tan, L. H. (1999). Transferable Licenses versus Nontransferable Licenses: What is the Difference? *International Economic Review*, **40**(3), 785–800.

Lapan, H. E. & Sikdar, S. (2019). Is Trade in Permits Good for the Environment? *Environmental & Resource Economics*, **72**(2), 501–10.

Lecuyer, O. & Quirion, P. (2013). Can Uncertainty Justify Overlapping Policy Instruments to Mitigate Emissions? *Ecological Economics*, **93**, 177 – 191.

MacKenzie, I. A. (2011). Tradable Permit Allocations and Sequential Choice. *Resource & Energy Economics*, **33**(1), 268–78.

Marchiori, C., Dietz, S. & Tavoni, A. (2017). Domestic Politics and the Formation of International Environmental Agreements. *Journal of Environmental Economics & Management*, **81**, 115–31.

Marschinski, R., Flachsland, C. & Jakob, M. (2012). Sectoral Linking of Carbon Markets: A Trade-Theory Analysis. *Resource & Energy Economics*, **34**(4).

Mehling, M. A., Metcalf, G. E. & Stavins, R. N. (2018). Linking Climate Policies to Advance Global Mitigation. *Science*, **359**, 997–8.

de Meza, D. & van der Ploeg, F. (1987). Production Flexibility as a Motive for Multinationality. *Journal of Industrial Economics*, **35**(3), 343–51.

Neuhoff, K., Schopp, A., Boyd, R., Stelmakh, K. & Vasa, A. (2012). *Banking of Surplus Emissions Allowances: Does the Volume Matter?* Working Paper 1196, DIW Berlin.

Newell, R. G. & Pizer, W. A. (2003). Regulating Stock Externalities under Uncertainty. *Journal of Environmental Economics & Management*, **45**(2), 416–32.

Newell, R. G. & Pizer, W. A. (2008). Indexed Regulation. *Journal of Environmental Economics & Management*, **56**(3), 221–33.

Newell, R. G., Pizer, W. A. & Zhang, J. (2005). Managing Permit Markets to Stabilize Prices. *Environmental & Resource Economics*, **31**(2), 133–57.

Newell, R. G. & Stavins, R. N. (2003). Cost Heterogeneity and the Potential Savings from Market-Based Policies. *Journal of Regulatory Economics*, **23**(1), 43–59.

Parsons, J. E. & Taschini, L. (2013). The Role of Stocks and Shocks Concepts in the Debate Over Price Versus Quantity. *Environmental & Resource Economics*, **55**(1), 71–86.

Pizer, W. A. (2002). Combining Price and Quantity Controls to Mitigate Global Climate Change. *Journal of Public Economics*, **85**(3), 409–34.

Pizer, W. A. & Prest, B. (2016). *Prices versus Quantities with Policy Updating*. Working Paper 22379, National Bureau of Economic Reasearch.

Pizer, W. A. & Yates, A. J. (2015). Terminating Links Between Emission Trading Programs. *Journal of Environmental Economics & Management*, **71**, 142–59.

Quemin, S. & de Perthuis, C. (2019). Transitional Restricted Linkage Between Emissions Trading Schemes. *Environmental & Resource Economics*, **74**(1), 1–32.

Quemin, S. & Trotignon, R. (2019). *Emissions Trading with Rolling Horizons*. Working Paper, Grantham Research Institute, London School of Economics.

Ranson, M. & Stavins, R. N. (2016). Linkage of Greenhouse Gas Emissions Trading Systems: Learning from Experience. *Climate Policy*, **16**(3), 284–300.

Roberts, M. J. & Spence, A. M. (1976). Effluent Charges and Licenses under Uncertainty. *Journal of Public Economics*, **5**(3), 193–208.

Rose, A., Wei, D., Miller, N., Vandyck, T. & Flachsland, C. (2018). Policy Brief – Achieving

Paris Climate Agreement Pledges: Alternative Designs for Linking Emissions Trading Systems. *Review of Environmental Economics & Policy*, **12**(1), 170–82.

Salant, S. W. (2016). What Ails the European Union's Emissions Trading System? *Journal of Environmental Economics & Management*, **80**, 6–19.

Samuelson, P. A. (1971). Stochastic Speculative Price. *Proceedings of the National Academy of Sciences*, **68**(2), 335–7.

Schennach, S. M. (2000). The Economics of Pollution Permit Banking in the Context of Title IV of the 1990 Clean Air Act Amendments. *Journal of Environmental Economics & Management*, **40**(3), 189–210.

Schopp, A., Acworth, W., Hauptmann, D. & Neuhoff, K. (2015). *Modelling a Market Stability Reserve in Carbon Markets*. Discussion Paper 1483, DIW Berlin.

Stevens, B. & Rose, A. (2002). A Dynamic Analysis of the Marketable Permits Approach to Global Warming Policy: A Comparison of Spatial and Temporal Flexibility. *Journal of Environmental Economics & Management*, **44**(1), 45–69.

Uribe, M. & Schmitt-Grohé, S. (2017). *Open Economy Macroeconomics*. Princeton University Press.

Weitzman, M. L. (1974). Prices *vs.* Quantities. *Review of Economic Studies*, **41**(4), 477–91.

World Bank (2019). *State and Trends of Carbon Pricing 2019*. Washington, DC: World Bank.

Yates, A. (2002). Decentralization in Pollution Permit Markets. *Journal of Public Economic Theory*, **4**(4), 641–60.

Yohe, G. W. (1978). Towards a General Comparison of Price Controls and Quantity Controls under Uncertainty. *Review of Economic Studies*, **45**(2), 229–38.

# Appendices

## A Analytical derivations and collected proofs

Throughout we will sometimes denote $\mathcal{G} = \{1, 2, \ldots, m\}$ for some $m \in [\![3; n]\!]$.

### A.1 Proof of Proposition 1 (effort- and risk-sharing gains)

Recalling the definition of $i$'s efficiency gains from $\mathcal{G}$-linkage in Equation (9), we have

$$\begin{aligned}\delta_{\mathcal{G},i} &= B_i(q_{\mathcal{G},i};\theta_i) - p_{\mathcal{G}}(q_{\mathcal{G},i} - \omega_i) - B_i(\omega_i;\theta_i) \\ &= (\beta_i + \theta_i - p_{\mathcal{G}})(q_{\mathcal{G},i} - \omega_i) - (q_{\mathcal{G},i}^2 - \omega_i^2)/(2\gamma_i) \\ &= q_{\mathcal{G},i}(q_{\mathcal{G},i} - \omega_i)/\gamma_i - (q_{\mathcal{G},i}^2 - \omega_i^2)/(2\gamma_i) \\ &= (q_{\mathcal{G},i} - \omega_i)^2/(2\gamma_i) = \gamma_i(p_i - p_{\mathcal{G}})^2/2,\end{aligned} \tag{A.1}$$

where the third and fifth equalities obtain via the first-order condition in Equation (6) and the net permit demand in Equation (8), respectively. Taking expectations and noting that $\mathbb{V}\{X\} = \mathbb{E}\{X^2\} - \mathbb{E}\{X\}^2$ for some random variable $X$ concludes the proof.

Next, we make explicit the sources of the efficiency gains generated in $\mathcal{G}$-linkage by characterizing the reduction in emissions control costs due to the link. Specifically, let $\tilde{a}_i = \tilde{q}_i - \omega_i > 0$ and $\Delta B_i$ respectively denote $i$'s domestic abatement level and the associated foregone benefits due to compliance with $i$'s binding cap under autarky, that is

$$\Delta B_i(\tilde{a}_i) = B_i(\tilde{q}_i;\theta_i) - B_i(\omega_i;\theta_i) = \tilde{a}_i^2/(2\gamma_i), \tag{A.2}$$

where the second equality follows from $\omega_i = \tilde{q}_i - \tilde{a}_i$ and $\tilde{q}_i = \gamma_i(\beta_i + \theta_i)$. By convexity of $\Delta B_i$, Jensen's inequality implies that an increase in uncertainty about laissez-faire emissions (and the corresponding cap stringency) raises the expected foregone benefits (or control costs) under autarky. Note that because $\theta_i$ is mean-zero, $\mathbb{E}\{\Delta B_i(\tilde{a}_i)\}$ can be decomposed as

$$\mathbb{E}\{\Delta B_i(\tilde{a}_i)\} = \Delta B_i(\bar{q}_i - \omega_i) + \mathbb{E}\{\Delta B_i(\tilde{q}_i - \bar{q}_i)\} = \gamma_i(\bar{p}_i^2 + \sigma_i^2)/2, \tag{A.3}$$

where the first term measures costs under certainty, which are proportional to $i$'s ambition level, and the second term captures the increase in costs due to uncertainty, which is proportional to the shock variance. By the same token, the aggregate expected control costs under

$\mathcal{G}$-linkage read

$$\sum_{i\in\mathcal{G}} \mathbb{E}\{\Delta B_i(\tilde{q}_i - q_{\mathcal{G},i})\} = \Gamma_{\mathcal{G}}(\bar{p}_{\mathcal{G}}^2 + \mathbb{V}\{\hat{\Theta}_{\mathcal{G}}\})/2. \tag{A.4}$$

Summing Equation (A.3) over $i \in \mathcal{G}$ gives the corresponding aggregate expected control costs under autarky. Note that $\sum_{i\in\mathcal{G}} \mathbb{E}\{\Delta B_i(\tilde{q}_i - q_{\mathcal{G},i})\} \le \sum_{i\in\mathcal{G}} \mathbb{E}\{\Delta B_i(\tilde{q}_i - \omega_i)\}$ as it jointly holds that $\Gamma_{\mathcal{G}}\bar{p}_{\mathcal{G}}^2 \le \sum_{i\in\mathcal{G}} \gamma_i \bar{p}_i^2$ and $\Gamma_{\mathcal{G}}\mathbb{V}\{\hat{\Theta}_{\mathcal{G}}\} \le \sum_{i\in\mathcal{G}} \gamma_i \sigma_i^2$.[29] In words, for given caps, linkage induces a cost-effective reduction in the group's expected control costs by (1) spreading the expected aggregate abatement effort in proportion to each member's technology and (2) improving the absorption of shocks within the linked system. Hence the effort- and risk-sharing gains.

## A.2 Proof of Proposition 2 (bilateral decomposition)

We first establish Equation (14). Substituting $p_{\mathcal{G}} = \sum_{i\in\mathcal{G}} \gamma_i p_i/\Gamma_{\mathcal{G}}$ into Equation (9) yields

$$\begin{aligned}\delta_{\mathcal{G},i} &= \gamma_i \Big(\sum_{j=1,j\neq i}^{m} \gamma_j (p_i - p_j)\Big)^2 \Big/ (2\Gamma_{\mathcal{G}}^2) \\ &= \gamma_i \sum_{j=1,j\neq i}^{m} \gamma_j \Big\{\gamma_j (p_i - p_j)^2 + 2 \sum_{k>j,k\neq i}^{m} \gamma_k (p_i - p_j)(p_i - p_k)\Big\} \Big/ (2\Gamma_{\mathcal{G}}^2).\end{aligned} \tag{A.5}$$

It is useful to note that the two following identities hold true

$$\begin{aligned}2(p_i - p_j)(p_i - p_k) &= (p_i - p_k + p_k - p_j)(p_i - p_k) + (p_i - p_j)(p_i - p_j + p_j - p_k) \\ &= (p_i - p_j)^2 + (p_i - p_k)^2 - (p_j - p_k)^2, \text{ and}\end{aligned} \tag{A.6}$$

$$\sum_{j=1}^{m} \sum_{k>j,k\neq i}^{m} \gamma_j \gamma_k \Big\{(p_i - p_j)^2 + (p_i - p_k)^2\Big\} = \sum_{j=1}^{m} \sum_{k=1,k\neq i,j}^{m} \gamma_j \gamma_k (p_i - p_j)^2. \tag{A.7}$$

Using these identities and rearranging the sums in Equation (A.5), we obtain that

$$\delta_{\mathcal{G},i} = \gamma_i \sum_{j=1}^{m} \gamma_j \Big\{(\Gamma_{\mathcal{G}} - \gamma_i)(p_i - p_j)^2 - \sum_{k>j,k\neq i}^{m} \gamma_k (p_j - p_k)^2\Big\} \Big/ (2\Gamma_{\mathcal{G}}^2). \tag{A.8}$$

Since the total $\{i,j\}$-linkage gains read $\Delta_{\{i,j\}} = \gamma_i\gamma_j(p_i - p_j)^2/(2\Gamma_{\{i,j\}})$ and $\Gamma_{\mathcal{G}\setminus\{i\}} = \Gamma_{\mathcal{G}} - \gamma_i$, Equation (A.8) coincides with Equation (14). Summing over all $i \in [\![1; m]\!]$ then gives

$$\Delta_{\mathcal{G}} = \sum_{i=1}^{m} \delta_{\mathcal{G},i} = \sum_{i=1}^{m} \Big\{ \sum_{j=1,j\neq i}^{m} \Big\{\Gamma_{\mathcal{G}\setminus\{i\}}\Gamma_{\{i,j\}}\Delta_{\{i,j\}} - \gamma_i \sum_{k>j,k\neq i}^{m} \Gamma_{\{j,k\}}\Delta_{\{j,k\}}\Big\}\Big\} \Big/ \Gamma_{\mathcal{G}}^2. \tag{A.9}$$

[29] The first and second inequalities hold strictly provided that there exists $(i,j) \in \mathcal{G}^2$ such that $\bar{p}_i \neq \bar{p}_j$ and respectively $\rho_{ij} < 1$ and/or $\sigma_i \neq \sigma_j$. See Appendix A.4 for a proof.

Regrouping terms by bilateral linkages, Equation (A.9) rewrites

$$\begin{aligned}\Delta_{\mathcal{G}} &= \sum_{1\le i<j\le m} \Big\{(\Gamma_{\mathcal{G}\backslash\{i\}} + \Gamma_{\mathcal{G}\backslash\{j\}})\Gamma_{\{i,j\}}\Delta_{\{i,j\}} - \sum_{k=1,k\neq i,j}^{m} \gamma_k \Gamma_{\{i,j\}}\Delta_{\{i,j\}}\Big\}\Big/\Gamma_{\mathcal{G}}^2 \\ &= \sum_{1\le i<j\le m} \Big\{(\Gamma_{\mathcal{G}\backslash\{i\}} + \Gamma_{\mathcal{G}\backslash\{j\}} - \Gamma_{\mathcal{G}\backslash\{i,j\}})\Gamma_{\{i,j\}}\Delta_{\{i,j\}}\Big\}\Big/\Gamma_{\mathcal{G}}^2 = \sum_{1\le i<j\le m} \Gamma_{\{i,j\}}\Delta_{\{i,j\}}\Big/\Gamma_{\mathcal{G}}.\end{aligned} \tag{A.10}$$

By symmetry, i.e. $\Delta_{\{i,j\}} = \Delta_{\{j,i\}}$, Equation (A.10) coincides with Equation (15).

As a side note, because variance is a symmetric bilinear operator, it holds that

$$\mathbb{V}\{\Delta_{\mathcal{G}}\} = \sum_{(i,j)\in\mathcal{G}\times\mathcal{G}} \Gamma_{\{i,j\}} \sum_{(k,l)\in\mathcal{G}\times\mathcal{G}} \Gamma_{\{k,l\}}\mathrm{Cov}\{\Delta_{\{i,j\}};\Delta_{\{k,l\}}\}\Big/(2\Gamma_{\mathcal{G}})^2. \tag{A.11}$$

Intuitively, although it is clear that $\mathcal{I} = \arg\max_{\mathcal{G}\subseteq\mathcal{I}} \mathbb{E}\{\Delta_{\mathcal{G}}\}$, there is no reason that forming larger groups reduces volatility of gains and a fortiori that $\mathcal{I} = \arg\min_{\mathcal{G}\subseteq\mathcal{I}} \mathbb{V}\{\Delta_{\mathcal{G}}\}$.

## A.3 Proof of Equation (17)

With $\mathcal{G}$ and $\mathcal{G}'$ in $\mathcal{I}$ such that $\mathcal{G}'\subset\mathcal{G}$ and with $\mathcal{G}'' = \mathcal{G}\backslash\mathcal{G}'$, expanding Equation (15) gives

$$\begin{aligned}\Delta_{\mathcal{G}} &= \Big(\sum_{(i,j)\in\mathcal{G}'\times\mathcal{G}'} \Gamma_{\{i,j\}}\Delta_{\{i,j\}} + \sum_{(i,j)\in\mathcal{G}''\times\mathcal{G}''} \Gamma_{\{i,j\}}\Delta_{\{i,j\}} + 2\sum_{(i,j)\in\mathcal{G}'\times\mathcal{G}''} \Gamma_{\{i,j\}}\Delta_{\{i,j\}}\Big)\Big/(2\Gamma_{\mathcal{G}}) \\ &= \Big(\Gamma_{\mathcal{G}'}\Delta_{\mathcal{G}'} + \Gamma_{\mathcal{G}''}\Delta_{\mathcal{G}''} + \sum_{(i,j)\in\mathcal{G}'\times\mathcal{G}''} \Gamma_{\{i,j\}}\Delta_{\{i,j\}}\Big)\Big/\Gamma_{\mathcal{G}}.\end{aligned} \tag{A.12}$$

The aggregate gains from linking $\mathcal{G}'$ and $\mathcal{G}''$ are $\Delta_{\{\mathcal{G}',\mathcal{G}''\}} = \Delta_{\mathcal{G}} - \Delta_{\mathcal{G}'} - \Delta_{\mathcal{G}''}$ so that

$$\begin{aligned}\Delta_{\{\mathcal{G}',\mathcal{G}''\}} &= \Big(\sum_{(i,j)\in\mathcal{G}'\times\mathcal{G}''} \Gamma_{\{i,j\}}\Delta_{\{i,j\}} + \Big(\Gamma_{\mathcal{G}'} - \Gamma_{\mathcal{G}}\Big)\Delta_{\mathcal{G}'} + \Big(\Gamma_{\mathcal{G}''} - \Gamma_{\mathcal{G}}\Big)\Delta_{\mathcal{G}''}\Big)\Big/\Gamma_{\mathcal{G}} \\ &= \Big(\sum_{(i,j)\in\mathcal{G}'\times\mathcal{G}''} \Gamma_{\{i,j\}}\Delta_{\{i,j\}} - \Gamma_{\mathcal{G}''}\Delta_{\mathcal{G}'} - \Gamma_{\mathcal{G}'}\Delta_{\mathcal{G}''}\Big)\Big/\Gamma_{\mathcal{G}}.\end{aligned} \tag{A.13}$$

By transposing Equation (13a) from two singletons to two groups, it holds that

$$\mathbb{E}\{\Delta_{\{\mathcal{G}',\mathcal{G}''\}}\} = \Gamma_{\mathcal{G}'}\Gamma_{\mathcal{G}''}\Big(\mathbb{V}\{p_{\mathcal{G}'}\} + \mathbb{V}\{p_{\mathcal{G}''}\} - 2\mathrm{Cov}\{p_{\mathcal{G}'};p_{\mathcal{G}''}\}\Big)\Big/(2\Gamma_{\mathcal{G}}) \geq 0. \tag{A.14}$$

## A.4 Proof of Proposition 3 (linking price properties)

For any $\mathcal{G}$ in $\mathcal{I}$, first note that price volatilities satisfy $\Gamma_{\mathcal{G}}\mathbb{V}\{p_{\mathcal{G}}\}^{1/2} \leq \sum_{i\in\mathcal{G}} \gamma_i \mathbb{V}\{p_i\}^{1/2}$ with a strict inequality provided that there exists $(i,j) \in \mathcal{G}^2$ such that $\rho_{ij} < 1$. Indeed,

$$\Gamma_{\mathcal{G}}^2\mathbb{V}\{p_{\mathcal{G}}\} = \sum_{(i,j)\in\mathcal{G}^2} \gamma_i\gamma_j \mathrm{Cov}\{p_i;p_j\} \leq \sum_{(i,j)\in\mathcal{G}^2} \gamma_i\gamma_j\sigma_i\sigma_j = \Big(\sum_{i\in\mathcal{G}} \gamma_i\mathbb{V}\{p_i\}^{1/2}\Big)^2. \tag{A.15}$$

Note that we have a similar inequality for price variances. Indeed, it jointly holds that

$$\Gamma_{\mathcal{G}}^2\mathbb{V}\{p_{\mathcal{G}}\} = \sum_{i=1}^m \gamma_i^2\sigma_i^2 + 2\sum_{1\leq i<j\leq m} \gamma_i\gamma_j\rho_{ij}\sigma_i\sigma_j, \text{ and} \tag{A.16a}$$

$$\Gamma_{\mathcal{G}}\sum_{j=1}^m \gamma_j\mathbb{V}\{p_j\} = \sum_{i=1}^m\sum_{j=1}^m \gamma_i\gamma_j\sigma_j^2 = \sum_{i=1}^m \gamma_i^2\sigma_i^2 + \sum_{1\leq i<j\leq m} \gamma_i\gamma_j(\sigma_i^2+\sigma_j^2). \tag{A.16b}$$

Then, $\Gamma_{\mathcal{G}}\mathbb{V}\{p_{\mathcal{G}}\} \leq \sum_{i\in\mathcal{G}} \gamma_i\mathbb{V}\{p_i\}$ follows since $\sigma_i^2+\sigma_j^2 \geq 2\rho_{ij}\sigma_i\sigma_j$ and observe that the inequality holds strictly when there exists $(i,j)\in\mathcal{G}^2$ such that $\rho_{ij}<1$ and/or $\sigma_i \neq \sigma_j$.

Formally, a partition $\mathcal{P}'$ of $\mathcal{I}$ is coarser (or finer) than partition $\mathcal{P}$ with $|\mathcal{P}| \geq 2$ and $d = |\mathcal{P}|-|\mathcal{P}'| \geq 1$ if there exists a sequence of partitions $(\mathcal{P}_i)_{i\in[\![0;d]\!]}$ such that $\mathcal{P}_0 = \mathcal{P}'$, $\mathcal{P}_d = \mathcal{P}$ and for all $i \in [\![1;d]\!]$ there exist $(\mathcal{G}_i', \mathcal{G}_i'') \in \mathcal{P}_i \times \mathcal{P}_i\backslash\{\mathcal{G}_i'\}$ such that $\mathcal{P}_{i-1} = \{\mathcal{G}_i' \cup \mathcal{G}_i''\} \cup \mathcal{P}_i\backslash\{\mathcal{G}_i', \mathcal{G}_i''\}$. It suffices to establish the rest of Statement $(a)$ for a unitary linkage since the proof extends to a more general case by transitivity over the relevant sequence of unitary linkages. Thus, let $\mathcal{P} = \{\mathcal{G}_1, \ldots, \mathcal{G}_z\}$ and assume w.l.o.g. that $\mathcal{P}' = \{\mathcal{G}_1\cup\mathcal{G}_2, \mathcal{G}_3, \ldots, \mathcal{G}_z\}$. Then, it holds that

$$\sum_{\mathcal{G}\in\mathcal{P}'} \Gamma_{\mathcal{G}}\mathbb{V}\{p_{\mathcal{G}}\}^{1/2} = \sum_{k=3}^z \Gamma_{\mathcal{G}_k}\mathbb{V}\{p_{\mathcal{G}_k}\}^{1/2} + \Big(\Gamma_{\mathcal{G}_1}^2\mathbb{V}\{p_{\mathcal{G}_1}\} + \Gamma_{\mathcal{G}_2}^2\mathbb{V}\{p_{\mathcal{G}_2}\} + 2\Gamma_{\mathcal{G}_1}\Gamma_{\mathcal{G}_2}\mathrm{Cov}\{p_{\mathcal{G}_1};p_{\mathcal{G}_2}\}\Big)^{1/2}. \tag{A.17}$$

The Cauchy-Schwarz inequality gives $|\mathrm{Cov}\{p_{\mathcal{G}_1};p_{\mathcal{G}_2}\}| \leq \mathbb{V}\{p_{\mathcal{G}_1}\}^{1/2}\mathbb{V}\{p_{\mathcal{G}_2}\}^{1/2}$ and concludes.

We now turn to Statement $(b)$. Note that it is sufficient to verify the claim on jurisdictional price variability as a result of linkage for bilateral links – the argument naturally extends to multilateral links. Then, by applying Equation (A.16a) to $\{i,j\}$-linkage it holds that

$$\mathbb{V}\{p_{\{i,j\}}\} = \Big(\gamma_i^2\mathbb{V}\{p_i\} + \gamma_j^2\mathbb{V}\{p_j\} + 2\rho_{ij}\gamma_i\gamma_j(\mathbb{V}\{p_i\}\mathbb{V}\{p_j\})^{1/2}\Big)\Big/\Gamma_{\mathcal{G}}^2. \tag{A.18}$$

Assume w.l.o.g. that jurisdiction $i$ is the less volatile jurisdiction, i.e. $\sigma_j \geq \sigma_i$. Then, $\{i,j\}$-linkage reduces price volatility in the high-volatility jurisdiction i.f.f. $\mathbb{V}\{p_j\} \geq \mathbb{V}\{p_{\{i,j\}}\}$, that is i.f.f.

$$\gamma_i(\sigma_j^2 - \sigma_i^2) + 2\gamma_j\sigma_j(\sigma_j - \rho_{ij}\sigma_i) \geq 0, \tag{A.19}$$

and unconditionally holds, i.e. for all $\gamma_i$, $\gamma_j$, $\sigma_j \geq \sigma_i$ and $\rho_{ij} \in [-1; 1]$. For the low-volatility jurisdiction, however, $\mathbb{V}\{p_i\} \geq \mathbb{V}\{p_{\{i,j\}}\}$ holds if and only if

$$\gamma_j(\sigma_i^2 - \sigma_j^2) + 2\gamma_i\sigma_i(\sigma_i - \rho_{ij}\sigma_j) \geq 0 \;\Leftrightarrow\; \frac{\gamma_j}{\gamma_i} \leq \frac{2\sigma_i(\sigma_i - \rho_{ij}\sigma_j)}{\sigma_j^2 - \sigma_i^2}. \tag{A.20}$$

For a given triple $(\sigma_i, \sigma_j, \rho_{ij})$, $\{i, j\}$-linkage effectively reduces volatility in the low-volatility jurisdiction provided that the high-volatility jurisdiction's $\gamma$ is relatively not too large.

Finally, to establish the claim on price convergence in probability, we let $\mathcal{G}$ be ordered such that $\gamma_1 \leq \ldots \leq \gamma_m$, and denote $\bar{\sigma} = \max_{i\in\mathcal{G}} \sigma_i$. Fix $\varepsilon > 0$. Then, it holds that

$$\begin{aligned}\mathbb{P}\Big(|\hat{\Theta}_{\mathcal{G}} - \mathbb{E}\{\hat{\Theta}_{\mathcal{G}}\}| > \varepsilon\Big) &\leq \mathbb{E}\Big\{(\hat{\Theta}_{\mathcal{G}} - \mathbb{E}\{\hat{\Theta}_{\mathcal{G}}\})^2\Big\}\Big/\varepsilon^2 = \mathbb{V}\{\hat{\Theta}_{\mathcal{G}}\}\Big/\varepsilon^2 \\ &= \sum_{i=1}^{m}\left\{\gamma_i^2\sigma_i^2 + \sum_{j=1}^{m}\rho_{ij}\gamma_i\gamma_i\sigma_i\sigma_j\right\}\Big/(\varepsilon\Gamma_{\mathcal{G}})^2 \\ &\leq \left(\frac{\gamma_m\bar{\sigma}}{\gamma_1\varepsilon}\right)^2\left[\frac{1}{m} + 1\right],\end{aligned} \tag{A.21}$$

where the first inequality is Chebyshev's inequality and the second follows by construction. Since $\gamma_m$ and $\bar{\sigma}$ are finite, only when the second term in the above bracket is nil (i.e. shocks are independent) does it hold that $p_{\mathcal{G}}$ converges in probability towards $\bar{p}_{\mathcal{G}}$ as $|\mathcal{G}|$ tends to infinity, that is $\lim_{m\to+\infty}\mathbb{P}\Big(|\hat{\Theta}_{\mathcal{G}} - \mathbb{E}\{\hat{\Theta}_{\mathcal{G}}\}| > \varepsilon\Big) = 0$, i.e. $\lim_{m\to+\infty}\mathbb{P}\Big(|\hat{\Theta}_{\mathcal{G}} - \mathbb{E}\{\hat{\Theta}_{\mathcal{G}}\}| \leq \varepsilon\Big) = 1$.

## A.5 A proof for the non-alignment of preferences

We prove the following claim for the non-alignment of linkage preferences:

*In the absence of inter-jurisdictional monetary transfers, jurisdictional linkage preferences are not aligned in the sense that*

*(a) $\mathcal{I}$-linkage may or may not be the most preferred linkage group for all jurisdictions in $\mathcal{I}$;*

*(b) any $\mathcal{G} \subset \mathcal{I}$ cannot be the most preferred linkage group for all jurisdictions in $\mathcal{G}$.*

Fix $\mathcal{G}' \subset \mathcal{I}$. Let $\mathcal{G} \supset \mathcal{G}'$ be a proper superset of $\mathcal{G}'$ and denote by $\mathcal{G}'' = \mathcal{G} \cap \mathcal{G}'$ the complement of $\mathcal{G}'$ in $\mathcal{G}$. By way of contradiction, assume that $\mathbb{E}\{\delta_{\mathcal{G}',i}\} \geq \mathbb{E}\{\delta_{\mathcal{G},i}\}$ holds for all $i \in \mathcal{G}'$, with at least one inequality holding strictly. By summation over $i \in \mathcal{G}'$

$$\sum_{i\in\mathcal{G}'}\mathbb{E}\{\delta_{\mathcal{G}',i}\} = \mathbb{E}\{\Delta_{\mathcal{G}'}\} > \sum_{i\in\mathcal{G}'}\mathbb{E}\{\delta_{\mathcal{G},i}\} = \mathbb{E}\{\Delta_{\mathcal{G}}\} - \sum_{i\in\mathcal{G}''}\mathbb{E}\{\delta_{\mathcal{G},i}\} \tag{A.22}$$

Recalling that $\Delta_{\{\mathcal{G}',\mathcal{G}''\}} = \Delta_{\mathcal{G}} - \Delta_{\mathcal{G}'} - \Delta_{\mathcal{G}''}$, Equation (A.22) imposes

$$\mathbb{E}\{\Delta_{\mathcal{G}''}\} + \mathbb{E}\{\Delta_{\{\mathcal{G}',\mathcal{G}''\}}\} - \sum_{i\in\mathcal{G}''} \mathbb{E}\{\delta_{\mathcal{G},i}\} < 0, \tag{A.23}$$

and contradicts superadditivity, which requires the above expression to be non-negative. That is, $\mathcal{G}'$ cannot be the most weakly preferred linkage coalition for all jurisdictions thereof.

## A.6 Proof of Proposition 4 (endogenous cap selection)

Let $D_i$ denote $i$'s damage function with $MD_i = \eta_i$ constant and positive. For any partition $\mathcal{P}$ of $\mathcal{I}$ we let $\Omega_{\mathcal{P}}^{-i} = \sum_{j\in\mathcal{I}\setminus\{i\}} \omega_{\mathcal{P},j}$ where $\omega_{\mathcal{P},j}$ is $j$'s cap given $\mathcal{P}$. Let also $\mathcal{A} = \{\{1\},\ldots,\{n\}\}$ denote complete autarky. The autarkic Cournot-Nash caps satisfy, for all $i \in \mathcal{I}$

$$\omega_{\mathcal{A},i} = \arg\max_{\omega>0} \mathbb{E}\Big\{ B_i(\omega;\theta_i) - D_i\left(\omega + \Omega_{\mathcal{A}}^{-i}\right) \Big\} = \gamma_i(\beta_i - \eta_i) > 0. \tag{A.24}$$

By identification with Equations (4) and (5) we find jurisdictional ambition parameters and expected autarky permit prices to be $\alpha_i = \beta_i - \eta_i \in (0;\beta_i)$ and $\bar{p}_i = \eta_i > 0$.

Jurisdictional regulators can anticipate linkage when selecting their caps. This situation is congruent with a two-stage game where regulators set caps at stage one and permit trading between linked markets occurs at stage two. We solve this game using backward induction and focus on subgame perfect Nash equilibria. Fix a partition $\mathcal{P}$ of $\mathcal{I}$. Crucially, because reaction functions are orthogonal, individual cap-setting decisions in any $\mathcal{G} \in \mathcal{P}$ will only be affected by the perspective of $\mathcal{G}$-linkage but not by what happens outside $\mathcal{G}$.

*Stage 2: Inter-jurisdictional permit trading and emissions choices.*

Take any $\mathcal{G} \in \mathcal{P}$. Given cap and realized shock profiles $(\omega_i)_{i\in\mathcal{G}}$ and $(\theta_i)_{i\in\mathcal{G}}$, Equations (7) and (6) respectively give the equilibrium permit price $p^*_{\mathcal{G}}$ and emission level in $i$ $q^*_{\mathcal{G},i}$

$$p^*_{\mathcal{G}}(\Omega_{\mathcal{G}};(\theta_i)_{i\in\mathcal{G}}) = \Big(\sum_{i\in\mathcal{G}} \gamma_i\beta_i - \Omega_{\mathcal{G}}\Big)\Big/\Gamma_{\mathcal{G}} + \hat{\Theta}_{\mathcal{G}}, \tag{A.25a}$$

$$\text{and } q^*_{\mathcal{G},i}(\Omega_{\mathcal{G}};(\theta_i)_{i\in\mathcal{G}}) = \gamma_i(\beta_i + \theta_i - p^*_{\mathcal{G}}(\Omega_{\mathcal{G}};(\theta_i)_{i\in\mathcal{G}})). \tag{A.25b}$$

We then obtain the intuitive comparative statics results: $\partial p^*_{\mathcal{G}}/\partial\Omega_{\mathcal{G}} = \partial p^*_{\mathcal{G}}/\partial\omega_i = -1/\Gamma_{\mathcal{G}} < 0$ and $\partial q^*_{\mathcal{G},i}/\partial\Omega_{\mathcal{G}} = \partial q^*_{\mathcal{G},i}/\partial\omega_i = \gamma_i/\Gamma_{\mathcal{G}} \in (0;1)$.

*Stage 1: Cournot-Nash domestic cap selection.*

Upon setting its cap, each regulator knows firms' optimal emission reactions and recognizes the implications of its decision on the expected permit price and its own net market position. The Cournot-Nash caps with strategic anticipation of $\mathcal{G}$-linkage $(\omega_{\mathcal{G},i})_{i\in\mathcal{G}}$ satisfy, for all $i$ in $\mathcal{G}$,

$$\begin{aligned}\omega_{\mathcal{G},i} = \arg\max_{\omega>0} \mathbb{E}\Big\{ & B_i\Big(q^*_{\mathcal{G},i}(\omega+\Omega^{-i}_{\mathcal{G}};(\theta_i)_{i\in\mathcal{G}});\theta_i\Big) - D_i(\omega+\Omega^{-i}_{\mathcal{G}}) \\ & + p^*_{\mathcal{G}}(\omega+\Omega^{-i}_{\mathcal{G}};(\theta_i)_{i\in\mathcal{G}})\Big(\omega - q^*_{\mathcal{G},i}(\omega+\Omega^{-i}_{\mathcal{G}};(\theta_i)_{i\in\mathcal{G}})\Big)\Big\},\end{aligned} \tag{A.26}$$

where the third term is the net proceeds from inter-jurisdictional permit trading. By stage-2 optimality, i.e. $\partial B_i(q^*_{\mathcal{G},i};\theta_i)/\partial q_i = p^*_{\mathcal{G}}$, the necessary first-order condition writes

$$\mathbb{E}\{p^*_{\mathcal{G}}\} - \eta_i = \mathbb{E}\Big\{\frac{\partial p^*_{\mathcal{G}}}{\partial \Omega_{\mathcal{G}}}(q^*_{\mathcal{G},i} - \omega_{\mathcal{G},i})\Big\} = (\omega_{\mathcal{G},i} - \mathbb{E}\{q^*_{\mathcal{G},i}\})/\Gamma_{\mathcal{G}}. \tag{A.27}$$

Summing over $i \in \mathcal{G}$ and by market closure, we obtain $\mathbb{E}\{p^*_{\mathcal{G}}\} = \langle\eta\rangle_{\mathcal{G}} = \sum_{i\in\mathcal{G}}\eta_i/|\mathcal{G}|$. It thus holds that $\omega_{\mathcal{G},i} - \mathbb{E}\{q^*_{\mathcal{G},i}\} = \Gamma_{\mathcal{G}}(\langle\eta\rangle_{\mathcal{G}} - \eta_i)$, i.e. jurisdiction $i$ is net selling under $\mathcal{G}$-linkage in expectation i.f.f. its marginal damage is lower than $\mathcal{G}$'s average. Since $\mathbb{E}\{q^*_{\mathcal{G},i}\} = \gamma_i(\beta_i - \langle\eta\rangle_{\mathcal{G}})$, we have

$$\omega_{\mathcal{G},i} = (\Gamma_{\mathcal{G}} - \gamma_i)(\langle\eta\rangle_{\mathcal{G}} - \eta_i) + \omega_{\mathcal{A},i} \geq \omega_{\mathcal{A},i} \;\Leftrightarrow\; \langle\eta\rangle_{\mathcal{G}} \geq \eta_i. \tag{A.28}$$

In aggregate, (anticipated) linkage leads to higher emissions relative to autarky i.f.f.

$$\sum_{i\in\mathcal{G}} \omega_{\mathcal{G},i} - \omega_{\mathcal{A},i} = \sum_{i\in\mathcal{G}} \gamma_i(\eta_i - \langle\eta\rangle_{\mathcal{G}}) \geq 0. \tag{A.29}$$

As in Helm (2003), inter-jurisdictional permit trading has an ambiguous effect on aggregate pollution, which depends on the distributions of the $\eta_i$'s and $\gamma_i$'s. Consider for instance the special case of marginal damages proportional to flexibilities, i.e. $\eta_i = \eta\gamma_i$ for all $i \in \mathcal{I}$. Then,

$$\begin{aligned}\sum_{i\in\mathcal{G}} \omega_{\mathcal{G},i} - \omega_{\mathcal{A},i} &= \eta\sum_{i\in\mathcal{G}}\gamma_i(\gamma_i - \langle\gamma\rangle_{\mathcal{G}}) = \eta\sum_{i\in\mathcal{G}}\gamma_i\Big((|\mathcal{G}|-1)\gamma_i - \sum_{j\neq i}\gamma_j\Big)\Big/|\mathcal{G}| \\ &= \eta\sum_{i\in\mathcal{G}}\sum_{j\neq i}\gamma_i(\gamma_i-\gamma_j)/|\mathcal{G}| = \eta\sum_{(i,j)\in\mathcal{G}^2}(\gamma_i-\gamma_j)^2/(2|\mathcal{G}|),\end{aligned}$$

which is always non-negative and positive provided that there exists $(i,j) \in \mathcal{G}^2$ such that $\gamma_i \neq \gamma_j$, but note that our result is reversed if we assume $\eta_i = \eta/\gamma_i$ for all $i \in \mathcal{I}$ instead.

Welfare gains from linkage accruing to jurisdiction $i$ belonging to any linkage group $\mathcal{G} \in \mathcal{P}$ amount to

$$B_i(q_{\mathcal{G},i};\theta_i) - B_i(\omega_{\mathcal{A},i};\theta_i) + p_{\mathcal{G}}(\omega_{\mathcal{G},i} - q_{\mathcal{G},i}) + D_i(\Omega_{\mathcal{A}}) - D_i(\Omega_{\mathcal{P}}). \tag{A.30}$$

By adding and subtracting $p_{\mathcal{G}}\omega_{\mathcal{A},i}$ it is convenient to decompose the expected welfare gains from linkage $\mathbb{E}\{\delta^*_{i,\mathcal{G}}\}$ into efficiency gains from inter-jurisdictional permit trading, strategic effect due to domestic cap selection in anticipation of linkage, and damage effect, that is

$$\begin{aligned}\mathbb{E}\{\delta^*_{i,\mathcal{G}}\} = \mathbb{E}\{&B_i(q_{\mathcal{G},i};\theta_i) - B_i(\omega_{\mathcal{A},i};\theta_i) + p_{\mathcal{G}}(\omega_{\mathcal{A},i} - q_{\mathcal{G},i})\} \\ &+ \mathbb{E}\{p_{\mathcal{G}}(\omega_{\mathcal{G},i} - \omega_{\mathcal{A},i})\} + \mathbb{E}\{D_i(\Omega_{\mathcal{A}}) - D_i(\Omega_{\mathcal{P}})\}.\end{aligned} \tag{A.31}$$

After standard computations, we find each of these components to be worth

$$\mathbb{E}\{B_i(q_{\mathcal{G},i};\theta_i) - B_i(\omega_{\mathcal{A},i};\theta_i) + p_{\mathcal{G}}(\omega_{\mathcal{A},i} - q_{\mathcal{G},i})\} = \gamma_i\Big((\eta_i - \langle\eta\rangle_{\mathcal{G}})^2 + \mathbb{V}\{\theta_i - \hat{\Theta}_{\mathcal{G}}\}\Big)\Big/2, \tag{A.32}$$

$$\mathbb{E}\{p_{\mathcal{G}}(\omega_{\mathcal{G},i} - \omega_{\mathcal{A},i})\} = \langle\eta\rangle_{\mathcal{G}}(\Gamma_{\mathcal{G}} - \gamma_i)(\langle\eta\rangle_{\mathcal{G}} - \eta_i), \tag{A.33}$$

$$\text{and } \mathbb{E}\{D_i(\Omega_{\mathcal{A}}) - D_i(\Omega_{\mathcal{P}})\} = \eta_i \sum_{\mathcal{G}\in\mathcal{P}} \sum_{j\in\mathcal{G}} \gamma_j(\langle\eta\rangle_{\mathcal{G}} - \eta_j). \tag{A.34}$$

## A.7 Proof of Proposition 5 (intertemporal trading)

We take the perspective of a group which might be degenerate, i.e. a single jurisdiction. The group-wide shock and benefit parameters (linear intercept and slope) obtain by horizontal summation of the individual marginal benefit schedules. In the following, we drop the group index for clarity and without loss of generality.

We adopt a dynamic programming approach and assume that time $t$ runs in $[\![1;T]\!]$ where $T$ is the date at which the problem effectively ends. Let $b_t \gtrless 0$ denote the volume of the permit bank at time $t$ (a negative bank corresponds to borrowing) with $b_0 = 0$ and $b_T \geq 0$. At each time $t$ the group emits $q_t = \omega_t + b_{t-1} - b_t$ and faces the recursive optimization problem

$$V_t(b_{t-1};\theta_{t-1}) = \max_{b_t} \Big[B_t(\omega_t + b_{t-1} - b_t;\theta_t) + \lambda\mathbb{E}_t\{V_{t+1}(b_t;\theta_t)\}\Big], \tag{A.35}$$

where $b_t$ is the control variable to simplify taking derivatives, and $\lambda$ denotes the discount factor. The first-order condition reads and the envelope theorem yields

$$-MB_t(q_t;\theta_t) + \lambda\mathbb{E}_t\Big\{\frac{\partial V_{t+1}}{\partial b_t}(b_t;\theta_t)\Big\} = 0, \tag{A.36a}$$

$$\frac{\partial V_t}{\partial b_{t-1}}(b_{t-1};\theta_{t-1}) = MB_t(q_t;\theta_t), \tag{A.36b}$$

so that we obtain the standard result that the discounted equilibrium marginal benefit (i.e. the discounted permit price) is a martingale via the no-arbitrage condition under uncertainty

$$MB_t(q_t;\theta_t) = \lambda \mathbb{E}_t\{MB_{t+1}(q_{t+1};\theta_{t+1})\}. \tag{A.37}$$

Let $h = T - t$ denote the number of future periods at time $t$. The period-$t$ equilibrium price with intertemporal trading $p_t^\star$ obtains through chaining the optimal law of motion across two adjacent periods in Equation (A.37) over the remaining $h$ periods

$$p_t^\star = MB_t(q_t;\theta_t) = \lambda \mathbb{E}_t\{MB_{t+1}(q_{t+1};\theta_{t+1})\} = \cdots = \lambda^h \mathbb{E}_t\{MB_{t+h}(q_{t+h};\theta_{t+h})\}, \tag{A.38}$$

where we have used the tower rule ($\mathbb{E}_t\{\mathbb{E}_{t+z}\{\cdot\}\} = \mathbb{E}_t\{\cdot\}$ for any $z \in [\![0;h]\!]$), together with overall market closure at the terminal date $b_T \geq 0$, or $\sum_{z=t}^T q_z \leq b_{t-1} + \sum_{z=t}^T \omega_t$. The period-$t$ expected price path satisfies Hotelling's rule.

To simplify computations and without loss of generality, we assume that $\beta_t = \beta_{t+1}$, $\gamma_t = \gamma_{t+1}$ and $\omega_t = \omega_{t+1}$ for all $t$. Solving Equation (A.38) with period-$T$ market clearing thus yields

$$\mathbb{E}_t\{p_{t+z}^\star\} = \lambda^{-z} p_t^\star \text{ for any } z \in [\![0;h]\!] \text{ with } p_t^\star = \Big((h+1)\bar{p} - b_{t-1}/\gamma + \theta_t \Phi\Big)\Big/\Lambda. \tag{A.39}$$

Setting $b_{t-1} = 0$ then gives Equation (24), and Equation (26) follows thanks to the linearity of both the group-wide shock in the individual shocks and the group's expected price in the expected autarky prices. Our determination of $p_t^\star$ (and below $\delta_{i,\mathcal{G},t}^\star$) can be extended to allow for time varying caps and benefit functions as well as heterogeneity in discounting.

Equation (27) obtains by computing $\delta_{i,\mathcal{G},t}^\star$ and taking expectation. Comparing Equations (10) and (27), we have $\mathbb{E}\{\delta_{i,\mathcal{G},t}^\star\} = \mathbb{E}\{\delta_{i,\mathcal{G},t}\}$ only when $h = 0$ or when $\varphi_i = 1$ for all $i \in \mathcal{G}$ and $\lambda = 1$. When $h \geq 1$, it is typically the case that $\mathbb{E}\{\delta_{i,\mathcal{G},t}^\star\} \neq \mathbb{E}\{\delta_{i,\mathcal{G},t}\}$ and their ordering will depend on the values of the jurisdiction-specific persistence parameters, the common discount factor and the length of the time horizon. We note that when $h \to \infty$ or $\lambda \to 0$, intertemporal permit trade attenuates the efficiency gains due to linking towards zero, i.e. $\mathbb{E}\{\delta_{i,\mathcal{G},t}^\star\} \to 0$. In particular, given an arbitrary $\lambda < 1$ there exists a threshold value of $h$ above which $\mathbb{E}\{\delta_{i,\mathcal{G},t}^\star\} < \mathbb{E}\{\delta_{i,\mathcal{G},t}\}$ holds unambiguously. When $h$ is small the ordering of $\mathbb{E}\{\delta_{i,\mathcal{G},t}^\star\}$ and $\mathbb{E}\{\delta_{i,\mathcal{G},t}\}$ depends on the complex interaction between $h$ and shock properties $\{\sigma_i, \rho_{ij}, \varphi_i\}_{i,j\in\mathcal{G}}$. We explore this analytically below in the case of a bilateral link.

The aggregate and jurisdictional risk-sharing gains due to $\{i,j\}$-linkage are proportional to

$$RS_\star = \Big(\Phi_i^2\sigma_i^2 + \Phi_j^2\sigma_j^2 - 2\rho_{ij}\Phi_i\Phi_j\sigma_i\sigma_i\Big)\Big/\Lambda^2 \text{ or } RS = \sigma_i^2 + \sigma_j^2 - 2\rho_{ij}\sigma_i\sigma_j \tag{A.40}$$

with or without intertemporal trading, respectively, where $\Phi_i = \sum_{z=0}^{h} \varphi_i^z$ and $\Lambda = \sum_{z=0}^{h} \lambda^{-z}$. Let prime denote the partial derivative w.r.t. $\varphi_i$, then we have

$$RS'_\star \geq 0 \;\Leftrightarrow\; \Phi'_i \Phi_i \sigma_i \geq \rho_{ij} \Phi'_i \Phi_j \sigma_j \tag{A.41}$$

since $(\Phi_i^2)' = 2\Phi'_i\Phi_i$ where $\Phi'_i = \sum_{z=1}^{h} z\varphi_i^{z-1}$. In general, $RS'_\star \gtreqless 0$ which crucially depends on the behavior of and the interaction between the series $\Phi_i$, $\Phi_j$, and $\Phi'_i$. The former, $\Phi_i$, is equal to $h+1$ when $\varphi_i = 1$; positive and increasing with $\varphi_i$ and $h$ when $\varphi_i \in (0;1)$; equal to 1 when $\varphi_i = 0$; non-monotonic in $\varphi_i$ and $h$ but non-negative for $\varphi_i \in (-1;0)$; alternating between 0 and 1 when $\varphi_i = -1$. Its partial derivative, $\Phi'_i$, is positive and increasing with $\varphi_i$ and $h$ when $\varphi_i > 0$; equal to 1 for all $h$ when $\varphi_i = 0$; non-monotonic in $\varphi_i$ and $h$ but non-negative for $\varphi_i \in [-0.5;0)$; of alternate sign (negative for $h \geq 3$ and odd) below some threshold $h$ value and then always positive when $\varphi_i \in (-1;-0.5)$; of alternate sign (negative for all $h$ odd) when $\varphi_i = -1$. In the two-period no-discount case $h = \lambda = 1$, Equation (A.41) simplifies to $RS'_\star \geq 0 \Leftrightarrow (1+\varphi_i)\sigma_i \geq \rho_{ij}(1+\varphi_j)\sigma_j$.

Similarly, it is not straightforward to compare $RS_\star$ and $RS$. Indeed,

$$RS_\star \geq RS \;\Leftrightarrow\; (\Phi_i^2 - \Lambda^2)\sigma_i^2 + (\Phi_j^2 - \Lambda^2)\sigma_j^2 - 2\rho_{ij}(\Phi_i\Phi_j - \Lambda^2)\sigma_i\sigma_j \geq 0, \tag{A.42}$$

The behavior of $\Phi_i^2$ is like that of $\Phi_i$. The series $\Lambda$ is equal to $h+1$ at most when $\lambda = 1$ for all $h$; increasing with $h$; decreasing with $\lambda$. By contrast, it is straightforward to show that allowing for intertemporal trading always reduces the effort-sharing gains from linkage by a fraction $((h+1)/\Lambda)^2$. In the two-period no-discount case $h = \lambda = 1$ with $\varphi_i = 1$, effort sharing is unaltered and Equation (A.42) simplifies to $RS_\star \geq RS \Leftrightarrow \rho_{ij}(3+\varphi_j)\sigma_i \geq \sigma_j$.

Finally, we clarify the mathematical statement in footnote 23 by specifying the behavior of the series $\Phi_i/\Lambda$. It is equal to 1 when $\varphi_i = 1$ and $\lambda = 1$; increasing with $\varphi_i$ and $\lambda$ and decreasing with $h$ when $\varphi_i \geq 0$; essentially increasing with $\varphi_i$ and $\lambda$ and decreasing with $h$ when $\varphi_i < 0$, although may be non-monotonic locally for small $h$ values and large $\lambda$ values.

# B Calibration methodology

This appendix describes the calibration of jurisdictional annual emission caps ($\omega_i$), baseline emissions ($\bar{q}_i$), volume-adjusted technologies ($\gamma_i$) and linear intercepts ($\beta_i$) based on proprietary data we obtained from Enerdata; and the calibration of price shock volatilities ($\sigma_i$), the pairwise correlations across jurisdictions ($\rho_{ij}$) and the AR(1) shock persistences ($\varphi_i$) based on IEA data on historical power sector emissions. In our quantitative illustration we focus on five jurisdictions with similar levels of development and which all use, or have considered using, both emissions trading and linking: Australia (AUS), Canada (CAN), the European Union (EUR), South Korea (KOR) and the United States (USA).

We obtained annual emissions caps and MACCs of the power sectors from Enerdata. First, Enerdata models emission caps consistent to three possible scenarios. The Ener-Brown scenario describes a world with durably low fossil fuel energy prices. The Ener-Blue scenario provides an outlook of energy systems based on the achievement of the 2030 targets defined in the NDCs as announced at COP 21. The Ener-Green scenario explores the implications of more stringent energy and climate policies to limit the global temperature increase at around 1.5-2$^o$C by the end of the century. We selected the scenario with annual emission caps consistent with the Paris INDCs (Ener-Blue scenario).

Second, Enerdata also generates MACCs and annual emission baselines using the Prospective Outlook on Long-term Energy Systems (POLES) model. MACCs are available for four time periods (2025, 2030, 2035 and 2040). We selected emission baselines and the MACCs available for 2030. Using these annual caps and MACCs, we compute the expected autarky permit prices, which range from 27.1 in AUS to 113.7\$/t$CO_2$ in Canada. All monetary quantities are expressed in constant 2005US\$. A linear interpolation of MACCs around domestic caps gives the linear intercept $\beta_i$ and the inverse of its slope $\gamma_i$, reported in Table 1.

The shock characteristics are calibrated using historical times series of $CO_2$ emissions from the jurisdictional power sectors. We obtain annual data covering 1972-2015 from the International Energy Agency. We denote observed emissions from jurisdiction $i$ in year $t$ by $e_{i,t}$. We identify historical emission levels with laissez-faire emissions, i.e. we assume that no or relatively lax regulations on $CO_2$ emissions were in place prior to 2015.

In Equation (2) laissez-faire emissions $\tilde{q}_i$ comprise a constant term, the baseline $\bar{q}_i = \gamma_i \beta_i$, and a variable term, $\tilde{q}_i - \bar{q}_i = \gamma_i \theta_i$. Assuming the latter is small enough relative to the former, we obtain the following linear Taylor approximation for the natural logarithm of laissez-faire emissions

$$\ln(\tilde{q}_i) \simeq \ln(\bar{q}_i) + (\tilde{q}_i - \bar{q}_i)/\bar{q}_i. \tag{B.1}$$

We associate the variable term in the above to the residual from the regression of $ln(e_{i,t})$ on time and the square of time. In other words, we use log-quadratic detrending to decompose $ln(e_{i,t})$ into trend and cyclical components (Uribe & Schmitt-Grohé, 2017). This is consistent with our interpretation of variations in marginal benefits of emissions as being driven by business cycles, TFP shocks, changes in the prices of factors of production, jurisdiction-specific events, weather fluctuations, etc.

Specifically, we denote the residuals from the regression $\epsilon_{i,t}$. To calibrate shock characteristics, we assume that $\{\epsilon_{i,t}\}$'s provide information about the distributions of the underlying shocks $\theta_i$'s. Then, given our modeling framework, $\epsilon_{i,t}$ is related to a draw from the distribution of $\theta_i$ such that

$$\epsilon_{i,t} = (\tilde{q}_i - \bar{q}_i)/\bar{q}_i = \theta_i/\beta_i. \tag{B.2}$$

Note that $\{\epsilon_{i,t}\}$'s are mean zero by construction. We compute the standard deviation of $\theta_i$ consistent with the model using

$$\sigma_i = \sigma(\beta_i \epsilon_{i,t}), \tag{B.3}$$

and the standard deviation of domestic laissez-faire power-sector emissions simply obtain by the rescaling $\gamma_i \sigma_i$. Table 3 below reports the standard deviations of autarky permit prices ($\sigma_i$) and normalized standard deviations of laissez-faire emissions ($\sigma(\epsilon_{i,t}) = \gamma_i \sigma_i / \bar{q}_i$). The table also includes the estimated persistence parameter $\varphi_i$ when an AR(1) model is fitted to $\{\epsilon_{i,t}\}$. We used the estimated $\varphi_i$'s to argue for the validity of the rule of thumb in Equation (28) discussed in the intertemporal trading extension in Section 4.2.

Table 3: Standard deviations of autarky prices ($\sigma_i$, 2005US\$/t$CO_2$), normalized standard deviations of laissez-faire emissions ($\sigma(\epsilon_{i,t}) = \gamma_i \sigma_i / \bar{q}_i$) and AR(1) shock persistences ($\varphi_i$)

| | EUR | AUS | USA | CAN | KOR |
|---|---|---|---|---|---|
| $\sigma_i$ | 31.4 | 11.8 | 21.9 | 56.3 | 48.4 |
| $\sigma(\epsilon_{i,t})$ | 0.049 | 0.054 | 0.058 | 0.131 | 0.112 |
| $\varphi_i$ | 0.79 | 0.71 | 0.82 | 0.89 | 0.67 |

Note that price shock variabilities are roughly such that $\bar{p}_i > 2\sigma_i$ and $\beta_i > \bar{p}_{\mathcal{G}} + 2\mathbb{V}\{\hat{\Theta}_{\mathcal{G}}\}^{1/2}$ for any jurisdiction $i$ and any possible group in our sample, i.e. zero-price and zero-emissions corners can safely be neglected.[30] Therefore, our focus on interior autarky and linking market equilibria is of negligible consequence for our analysis of linkage gains.

Finally, we calibrate pairwise correlation between shocks in $i$ and $j$ using

$$\rho_{ij} = \text{Corr}(\beta_i \epsilon_{i,t}, \beta_j \epsilon_{j,t}). \tag{B.4}$$

[30] Note that a sufficient condition for the second type of inequalities to hold is $\beta_i > \bar{p}_i + 2\sigma_i$ for all $i$.

and note that the $\rho_{ij}$'s – reported in Table 2 – can be positive, negative or approximately zero. We also note that this large variation in inter-jurisdictional correlation is to be expected.

To see why note that emissions of jurisdictions whose economies are tightly interconnected through trade and financial flows will likely move together, especially if jurisdictions' emissions are procyclical. If the economic links between jurisdictions are weak and/or they are geographically distant, one would expect a low level of correlation. Finally, if a jurisdiction's business cycles are negatively correlated with others, also observing negative correlations in emissions fluctuations would not be surprising. These conjectures are consistent with empirical studies such as Calderón et al. (2007) which provides evidence on international business cycle synchronization and trade intensity, and Doda (2014) which analyzes the business cycle properties of emissions. Finally, Burtraw et al. (2013) suggest that demand for permits may be negatively correlated over space due to exogenous weather shocks.

We highlight the following three points regarding our calibration strategy and results. First, we assume that the pair characteristics are not affected by the recent introduction of climate change policies. Some emitters in some of the jurisdictions in our sample are regulated under these policies. We argue that any possible effect would be limited because these policies have not been particularly stringent, affect only a portion of the jurisdiction's emissions, and do so only in the last few years of our sample. Second, we use the log quadratic filter to decompose the observed emissions series into its trend and cyclical components. Not surprisingly, the calibrated shock characteristics are altered quantitatively when we alternatively use the band pass filter recommended by Baxter & King (1999), the random walk band pass filter recommended by Christiano & Fitzgerald (2003) or the Hodrick-Prescott filter as detrending procedures. However, our conclusions are similar qualitatively so we restrict our attention to the simple and transparent log quadratic detrending. Third, we take the calibrated $\rho_{ij}$'s at face value in our computations, rather than setting insignificant correlations to zero, which does not alter the results in a meaningful way.