DISCUSSION PAPER SERIES

No. 10671

JOINT DESIGN OF EMISSION TAX AND TRADING SYSTEMS

Bernard Caillaud and Gabrielle Demange

PUBLIC ECONOMICS

Centre for Economic Policy Research

ISSN 0265-8003

JOINT DESIGN OF EMISSION TAX AND TRADING SYSTEMS

Bernard Caillaud and Gabrielle Demange

Discussion Paper No. 10671
June 2015
Submitted 11 June 2015

Centre for Economic Policy Research
77 Bastwick Street, London EC1V 3PZ, UK
Tel: (44 20) 7183 8801
www.cepr.org

JOINT DESIGN OF EMISSION TAX AND TRADING SYSTEMS[†]

Abstract

This paper analyzes the joint design of fiscal and cap-and-trade instruments in climate policies under uncertainty. Whether the optimal mechanism is a mixed policy (with some firms subject to a tax and others to a cap-and-trade) or a uniform one (with all firms subject to the same instrument) depends on parameters reflecting preferences, production, and, most importantly, the stochastic structure of the shocks affecting the economy. This framework is then used to address the issue of the non-cooperative design of ETS in various areas worldwide and to characterize the resulting inefficiency and excess in emission. We provide a strong Pareto argument in favor of merging ETS of different regions in the world and evaluate the welfare gains in each region.

JEL Classification: D62, H23 and Q54
Keywords: cap-and-trade mechanisms, climate policies and tax

Bernard Caillaud caillaud@pse.ens.fr
Paris School of Economics – Ecole des Ponts ParisTech and CEPR

Gabrielle Demange demange@pse.ens.fr
Paris School of Economics – EHESS and CEPR

[†] This research is an outgrowth of a CEPREMAP 'Grand Project' that the authors developed in collaboration with J. Pouyet, who should receive some of the credit for this paper. The authors are grateful to CEPREMAP for its financial support. They have benefited from comments from invited researchers at PSE and participants in the Conference 'Instruments to Curb Global Warming: Recent Developments,' co-organized with the CDC-Climat Recherche in Paris, the 5th Euro-African Conference in Finance and Economics in Agadir, and the World Congress of Environmental and Resource Economists in Istanbul, PEES in Paris.

1 Introduction

It is widely acknowledged today that the post-Kyoto era, which mainly relies on the negotiation process with respect to the Kyoto protocol second commitment period 2012-2020, is plagued with uncertainty and suffers from the absence of clear agreements and of a wide enough consensus among countries. Indeed, the Kyoto protocol first commitment period itself may be characterized by a general failure of its objectives. On the one hand, GHG emission has not started to decrease at the end of the first commitment period and, according to experts, global warming at an acceptable 1^oC within a century is not an attainable target any longer. On the other hand, during the first commitment period, neither an acceptable and consensual framework for the design of climate policies nor a global architecture have emerged that would pave the way for an easy-to-reach agreement among countries of the core (basically Annex I countries minus Canada) in order to convince reluctant parties to the Treaty to enter a global mechanism.

There are of course many explanations to these failures, among which the fact that the US never ratified the Treaty, the fact that developing countries were subject to only light commitment, and the cost of necessary measures today for uncertain benefits remote in the future. Also, there is no consensus on how Kyoto commitment should be implemented at the decentralized levels of countries, or of areas such as the EU. Indeed, even among economists, the discussion goes on about the relative merits of fiscal instruments compared to cap-and-trade mechanisms.[1] The cap-and-trade approach seems more successful recently as various forms of Emission Trading Systems (ETS) have been adopted in a few areas across the world,[2] even in the non-ratifying US or in withdrawing Canada.[3] Yet, it is striking that these systems have not been elaborated more cooperatively and that their mechanisms exhibit many differences.[4]

Of course, the emergence of so many ETS has raised the issue of linking various ETS as, from a standard economic point of view, linking two markets trading the same good is welfare-improving. Though, there are difficulties and resistance. A difficulty is well known: A Pareto improvement may call for transfers, which are not easy to implement. Also, non-economic features, such as the reliability of the trading system and of enforcement mechanisms are put forward by the EU (EU Report, 2008). Finally, difficulties arise due to the differences in the design of the existing or planned ETS. For example, the Australian White Paper (2008) was advocating a cap on the price of the permits for

[1]See Guesnerie (2010) for a survey.

[2]Early starters are the Australian NSW (2003) and the EU ETS (2005). The EU ETS has now integrated Norway domestic emission trading, which started in 2005 too, and the UK ETS that started as early as 2002. The Swiss ETS ran 2008-2012, the Japan ETS for the Tokyo area started in 2010. The New Zealand ETS started in 2008.

[3]The Regional Greenhouse Gas Initiative (RGGI) started in 2009 and it caps emissions from power generation in ten north-eastern US states. Emissions trading in California and the Western Climate Initiative (WCI, a collective ETS agreed between 11 US states and Canadian provinces) are only a couple of years old.

[4]For a description of the existing or planned ETS as of 2013 see Talberg - Swoboda (2013).

the planned Australian ETS,[5] while the EU ETS has no such cap. The total amount of permits allocated to the firms submitted to an ETS and how this amount should evolve over time differ across areas. Also ETS may cover different sets of industries: e.g. the planned coverage of the Australian ETS was larger than the current EU ETS, which does not include transportation nor forestry. Presumably, all these differences stem from differences in countries' characteristics, such as their "preferences", their production profiles, their appreciation of the impact of an ETS, and the influence of their lobbies.

As these discussions suggest, a theory of the optimal design of the architecture of climate regulation and the precise determination of climate policies is missing. This, however, seems to be a necessary first step in order to discuss the relative merits of various designs of ETS or even to evaluate the opportunity of linking ETS of different areas together. In this paper, we propose such a theoretical foundation: we propose a simple normative model that enables us to provide a meaningful discussion about the welfare-maximizing design of both fiscal and cap-and-trade instruments under the assumption that the economy is affected by shocks. We analyze how the optimal regulation depends on some parameters reflecting preferences, production and the structure of the shocks affecting the economy. We then use this benchmark to evaluate the losses due to non-coordination in the design of climate regulation and climate policies among various local regulators around the world. This leads us to analyze the efficiency gains that can be expected from linking different ETS with possibly different coverage.

Our analysis relies on a static model that extends Weitzman (1974) by allowing for the possibility of the double control mechanism. The mechanism specifies which firms are submitted to a tax and which ones to an ETS, what we call the *scope* of the regulation, as well as the associated tax level and quota allocated to the ETS, what we call the *policy*. The mechanism is decided ex ante, before the realization of the shocks that affect the firms in the economy. This captures the fact that the regulatory framework cannot be contingent on the realization of shocks that hit constantly the economy. Yet, the firms' reaction to shocks should be taken into account when designing a regulatory framework. We characterize optimal (or equilibrium) policies for any scope and we explain how the stochastic structure of the shocks influences the optimal design of the scope of the regulatory framework. In particular, we analyze when it is preferable to adopt a uniform system, subjecting all firms to either a cap-and-trade mechanism or a tax on their emissions, or a mixed system in which some firms are regulated through a cap-and-trade mechanism and the others through a tax.

The basic forces at play are the following. If climate regulation could be made contingent on shocks, i.e. a first best scenario, abatement efforts should be determined so as to equalize marginal abatement cost across firms with the social marginal benefit of abatement: part of the aggregate shocks should then be absorbed through abatement at the

[5]This ETS was supposed to start in 2015 but all climate legislation has been repealed in July 2014 due to a political swing.

firms' level and the coefficient of absorption would be larger, the less steep the aggregate marginal abatement cost curve, and the steeper the marginal abatement benefit curve.

In a regulatory framework characterized by a ETS sector and a taxed sector, the ETS sector absorbs all the shocks that impact it, which induces fluctuations in the corresponding marginal cost of abatement, while the taxed sector has a constant marginal cost of abatement but generates a random volume of emissions that reflects entirely the shocks that impact it. The optimal tax rate and ETS quota are determined so as to replicate the first best optimum in expected terms. Expected marginal abatement costs, that is the tax rate and the expected price on the ETS market, are equalized to their first best value; and expected net emissions are equalized to their first best optimal value as well.

So, the precise definition of the scope of the regulatory framework, i.e. of the industries to be included in the ETS and of those to be taxed, only affects social welfare through the fluctuations due to the shocks. The optimal scope should then be designed so as to replicate as closely as possible the emissions fluctuations corresponding to the first best allocation, given that all fluctuations in emissions are generated by firms subject to the tax. A uniform ETS system, in which all firms are subject to the ETS regulation, eradicates all fluctuations while a uniform tax system induces all shocks to be passed on in emissions fluctuations. Comparing both systems amounts to assessing the relative slopes of the marginal abatement cost and benefit curves, as in Weitzman (1974). Improving on either system requires to analyze mixed systems, with a non-degenerate ETS sector and a non-degenerate taxed sector, and to calibrate the taxed sector so that the shocks that affect it are sufficiently correlated with the partially dampened aggregate shocks as required in the first best. But doing so creates a wedge between the marginal abatement cost of ETS firms and taxed firms, hence a social loss due to the misallocation of abatement efforts across firms. The optimal scope optimally balances these effects.

We then use this framework to address the issue of the non-cooperative design of ETS in various areas worldwide. We consider a world consisting of several areas, in which each area uses a double control mechanism. The non-cooperative outcome is compared to the first best emission levels for the global economy and the corresponding inefficiency that results, i.e. excess in emissions worldwide, is precisely analyzed. Moreover, we analyze the proposal of linking ETS by specifying a specific form of linking, that we call ETS merging. We provide a strong Pareto argument in favor of merging ETS: such a move benefits both areas, even without implementing transfers across them or changing the sovereign decisions with respect to the fiscal instruments, and we precisely characterize these benefits for each area. Within each area, there are possible losers -for example the firms if the ETS price in the area is usually lower without the merger than in other ETS, but they may be compensated within the area by adequate transfers due to the extra resources collected by the government.

Our paper contributes to the long literature, pioneered by Weitzman (1974), on the use of price or quantity instruments in a framework characterized by uncertainty and asymmetric information about the shocks between the central authority and the economic agents. Within a framework of uncertainty on compliance costs, various ways of combining price and quantity instruments have been shown to provide welfare improvements: a three-part tariff (Roberts and Spence 1976), indexed or hybrid instruments allowing a variable quota (Pizer 2002 and Newell and Pizer 2008), joint use of price and quantity regulation in the context of multiple pollutants depending on their degree of substitutability or complementarity (Ambec and Coria 2013). Specifically, our paper is related to Mandell (2008) who analyzes the possibility of a mixed system in which some firms are subject to a price mechanism while others are subject to a quantity mechanism. Compared to our paper, Mandell (2008) focuses on the optimal scope in a more restricted framework in which there is a single common shock affecting all firms. He shows that a mixed system might be superior to a single uniform system. The first part of our analysis can be viewed as providing a generalization of this argument to general stochastic structures. Independently of our paper, Carlen and Hernandez (2013) consider a general structure of shocks (with imperfect correlation) but they restrict their analysis to two firms: in this setting, they obtain a simple version of our result on the minimal degree of positive correlation that is necessary to obtain an optimal mixed system.[6]

There is ample evidence that the volatility of the carbon price is large and various studies have analyzed empirically its determinants. In particular, Chevallier (2011) analyzes both the impact of industrial production and energy prices on the carbon market and confirms that both have an impact.[7] This justifies our modeling choice, namely that the shocks to the economy play a crucial role in the determination of the optimal scope as they drive the volatility of both the ETS price and the emission of the firms submitted to the emissions tax.

The second part of our paper relates to the more recent but very active literature on the linkage of different climate regulation settings, most notably the linkage of different ETS. The linkage of ETS is becoming the hotest topic in the debate about the evolution of the architecture of climate policies.[8] The linkage of carbon policies raises difficulties because different countries (or areas) have different objectives / preferences or different targets (see e.g. Metcalf and Weisbachy (2012)). Most of this literature assumes away uncertainty and starts with an initial carbon policies in each country that possibly reflect the country's objective but are globally inefficient. ETS merging then can improve global welfare as it induces cost equalization among participating firms. But differences in objectives also lead

[6]In addition, they investigate the possibility of indexing the tax rate on the realized value of the ETS price: this possibility of indexing the tax rate on the realization of some uncertainty is potentially efficiency enhancing.

[7]More precisely, the author considers a Markov-switching VAR model with two states that is able to reproduce the boom - bust business cycle. Industrial production is found to impact positively (resp. negatively) the carbon market during periods of economic expansion (resp. recession), and the energy prices impact the Markov-switching model.

[8]As reflected by the common thread in the contributions of the special issue of *Climate Policy* (Volume 9, Issue 4, 2009); see also Ranson and Stavins (2012).

countries to incorporate different cost containment or price control mechanisms in their local ETS and to take differently into account offset mechanisms such as CDM. These differences constitute major obstacles to the merging of ETS. For example, Jotzo and Betz (2009) focus on the difficulties raised by some important specifications of the Australian project, i.e. a price cap with unlimited access to international CDM. Linking the EU ETS scheme with such a scheme would have a large impact by effectively introducing a price cap for the global system and bypassing the European constraint on the use of offset mechanisms.[9]

Compared to this literature, we adopt a simplified setting in which we rule out CDM and other offset mechanisms, and we consider that ETS markets are perfectly competitive with no control mechanisms. We, however, endogenize the (centralized and decentralized) determination of carbon policies by explicitly formalizing the role of uncertainty as the fundamental determinant in their design. Wood *et al.* (2013) and Heindl - Wood - Jotzo (2014) rely also on a stochastic model to assess various ways of linking carbon policies between two possibly asymmetric countries. Both papers take the situation of ETS merging as a benchmark. Wood *et al.* (2013) compares it with a situation in which one country is under a tax system and not a ETS and countries trade allowances, while Heindl - Wood - Jotzo (2014) compares it to the situation in which one country imposes an additional tax on its constituent firms, on top of the joint ETS requirement. By contrast, our paper provides an endogenous characterization of the mixed systems adopted in each country under separate regulation, taking explicitly into account the impact of shocks on welfare; then we focus on the ETS merging scenario as compared to the situation of separate systems and we provide a quantitative assessment of the ensuing welfare gains in each country.

The plan of the paper is as follows. Section 2 provides a normative analysis of climate regulation from a worldwide perspective. Section 3 turns to the choices of several areas with two focuses: we first analyze the non-cooperative choices of regulatory instruments, and then we consider the incentives to merge ETS. Technical results are proved in two appendices, gathering results corresponding to the two main sections.

2 A normative analysis of climate policies

We consider a global economy in which the production process generates a stochastic volume of emissions of a pollutant. Consumers care about the aggregate emissions volume. Firms may reduce emissions through costly abatement if they have incentives to do so.

[9] Tuerk *et al.* (2009) and Sterk and Kruger (2009) argue similarly that very few direct or bilateral merging of ETS would be viable in the short run. They also discuss why an indirect linkage emerges among carbon markets, through the recognition of CDM and other crediting mechanisms, and how it may help improve global efficiency. Anger (2008), in a similar vein, shows that significant and beneficial effects can arise from opening the Kyoto system, so far restricted to governments, to ETS firms, thereby effectively creating a world market. Flachsland *et al.* (2009) points out ETS market failures as well as strategic manipulation of national carbon policies as additional obstacles to viable merging of ETS.

We restrict the analysis to two incentive instruments: an emission trading system (ETS) and a tax. Both instruments can be used simultaneously, with some activities subject to a tax and the others to an ETS.

The main questions we investigate in this section are: How to determine the firms covered by an ETS and those subject to a tax ? How to determine the quota of emissions allocated to the ETS firms and the tax rate imposed on non-ETS firms ? Our answers provide a normative approach to the design of climate policies from a worldwide perspective.

2.1 The model

We make two main modeling assumptions. First, we suppose there is separability in terms of costs and welfare between the markets for the goods and the emissions volume: abatement decisions have no impact on the goods' equilibrium prices and traded quantities. Second, we introduce uncertainty in the form of shocks on gross (pre-abatement) levels of emissions and we rule out shocks on the abatement technology itself. Both assumptions are restrictive but standard from the extant literature. We also make a third milder assumption: all cost and surplus functions are quadratic with respect to emissions volumes. This assumption can be viewed as an approximation of more general functional forms.

Firms. There are n firms in the economy. Each firm i, $i \in N \equiv \{1, 2, ..., n\}$, emits a volume of pollutant through its production activities when it does not make any abatement effort. Let $z_i + \epsilon_i \geq 0$ denote this volume. z_i is assumed to be common knowledge in the economy; ϵ_i is known by firm i at the time it decides how much to abate.

Firm i has access to an abatement technology with linear marginal cost:[10] abating $a_i \geq 0$ costs $\frac{b_i}{2} a_i^2$ and reduces firm i's emissions down to a volume of net emissions $x_i = z_i + \epsilon_i - a_i$. The smaller b_i, the more elastic the firm's abatement decisions to a variation in the unit price of emissions: $1/b_i$ can thus be viewed as a measure of firm i's flexibility in abating.

At the time of the design of the regulatory instruments, the values of each ϵ_i are unknown, perceived as random. The random variable $\widetilde{\epsilon}_i$ is referred to as firm i's shock. W.l.o.g. we suppose that $\widetilde{\epsilon}_i$ has zero mean so that z_i is firm i's average emission volume. The structure of all shocks $(\widetilde{\epsilon}_i)$ in the economy plays an important role in the analysis and we provide below examples of the factors that shape this stochastic structure.

- An increase in the emission by-product for a given input use in firm i, e.g. because of a deterioration (aging) in its production process or technology, induces a direct increase in the firm's gross emissions volume and consequently an increase in the marginal abatement cost to achieve a given level of net emissions.

[10] A linear term of the form $\mu_i a_i$ can be added without change in the analysis, up to a translation in the expressions for the emission volumes, prices and taxes.

- A decrease in the prices of carbon-intensive inputs, e.g. a decrease in the oil prices, leads to an increase in the use of these inputs; this induces an increase in the gross emissions volumes of all firms using such an input and ultimately to an increase in the marginal abatement cost of these firms for a given level of net emissions; moreover these increases are correlated across all firms using this input.

- An increase in the demand for the goods in some industry increases the gross level of emissions through an increase in production, hence ultimately increases the marginal abatement cost for given net emissions for any firm in this industry; these effects are strongly correlated across the firms within the industry. For the same reason, a macroeconomic shock that affects all industries induces correlation across industries.

Firms are submitted to an emission tax or to an ETS, as will be described below. For the moment, let τ denote the unit cost firm i is facing. Firm i choosing to emit x_i obtains net profits equal to:

$$\Pi_i = \xi_i - \frac{b_i}{2}\left(z_i + \epsilon_i - x_i\right)^2 - \tau x_i, \tag{1}$$

in which ξ_i summarizes the net profits on the goods' markets, possibly affected by some shocks, but independent of the net emissions volume.[11]

Given a set S of firms, $S \subseteq N$, let us define:

$$z_S = \sum_{i\in S} z_i, \epsilon_S = \sum_{i\in S} \epsilon_i \text{ and } \frac{1}{b_S} = \sum_{i\in S} \frac{1}{b_i}, \tag{2}$$

with the convention that $1/b_S = 0$ is S is empty.

Let us interpret these expressions. z_S is the total expected gross emissions volume for a group S of firms, e.g. a sector or an industry, and ϵ_S is the total shock in gross emissions of this group. Thus $z_S + \epsilon_S$ corresponds to the gross emissions level of group S. Aggregating the abatement technologies of all firms in the group S, the abatement cost at the group's level is quadratic given by $\frac{b_S}{2}a^2$; the reason is that an abatement of a units by the firms in S is efficiently obtained by assigning shares to the firms in proportion of their flexibility, i.e. by assigning ab_S/b_i to i. This results in the flexibility $1/b_S$ at the group level. When all the b_i are equal to b, b_S equals $b/|S|$ because a marginal increase in abatement can be equally distributed among all the firms within the group and therefore limits the impact of decreasing returns in abatement at the level of each firm.

Consumers surplus and social welfare. Let $X \equiv \sum_{i\in N} x_i$ denote the aggregate level of emissions in the economy. Consumers' loss due to emissions only depends on total emissions level X and is also assumed to be quadratic. In terms of consumers' surplus,

[11] The net profits on the goods markets are thus independent of the environmental policy. This is a restrictive assumption that can be relaxed as we discuss in Section 2.4.

this gives:

$$S = \lambda - \nu X - \frac{A}{2} X^2 \tag{3}$$

where λ summarizes the surplus on the goods markets [12]. This form for the consumers' surplus corresponds to a linearly increasing (social) marginal abatement benefit, equal to: $\nu + AX$. A measures the slope of the marginal abatement benefit curve: it may be perceived as large if e.g. the current volume of emissions is such that catastrophic climatic consequences would follow a small increase in emissions given the current situation (threshold effect).

Finally, the revenues R from the emission tax or the sales of permits are collected by a governmental agency. Assuming no cost of public funds, social welfare is given by $W = S + R + \sum_{i \in N} \Pi_i$. Transfers within the economy are socially neutral. From (1) and (3), the expected welfare is (up to an additive term independent of the emission levels):

$$\mathbb{E}\left\{ -\sum_{i \in N} \frac{b_i}{2} \left(z_i + \epsilon_i - x_i\right)^2 - \nu \left(\sum_{i \in N} x_i\right) - \frac{A}{2} \left(\sum_{i \in N} x_i\right)^2 \right\}. \tag{4}$$

Optimality is defined with respect to this social welfare criterion.

The designer aims at maximizing this expression, taking into account the reaction of the firms and their knowledge of the shocks. We focus on situations in which there are restrictions on the regulatory tools that can be used : some firms are regulated by a cap-and-trade system and the others by an emissions tax, as we describe now.[13]

Scope and policy. Let N be partitioned in two subsets, T and Q with $N = T \cup Q$. Firms in T are subject to the same tax rate t. Firms in Q are subject to a cap-and-trade, in which an amount $\overline{X}_Q$ of emissions is allowed and a (perfectly competitive) resale market operates so as to reallocate emission allowances across firms. (T, Q) is called the *scope* of the system and, given the scope, the *policy* consists in $(t, \overline{X}_Q)$. The system is said to be *uniform* when $T = N$ (all firms are submitted to a tax) or $Q = N$ (all firms are submitted to the ETS).

We use the term "firm" for simplicity. In practice, the assignment to an ETS is made at the level of the activity of a plant, so that a company may have only some plants submitted to an ETS. Furthermore, the assignment is not discretionary in the sense that all plants performing the same activity should be treated the same way, i.e. all assigned to the ETS or none. Our analysis is therefore better interpreted as applying to activities instead of firms; in that interpretation, the abatement cost of i refers to the average cost of the plants performing activity i.

[12] λ is an exogenous random variable independent of the shocks ϵ_i; the same remark as in footnote 11 applies here.

[13] We consider a single tax level and a single trading system. This is not a restrictive assumption as will be clear later on.

The scope and the policy are implemented before the shocks are realized. Then, after the scope and the policy are determined, uncertainty resolves and firms react by choosing emission levels so as to optimize their net profits and choose their positions on the ETS. The aim of this section is to investigate the scope and policy that maximize welfare. To understand the various inefficiencies associated with the scope design, we first consider the *first best* optimal allocation.

First best allocation. The *first best* optimal allocation maximizes the social welfare without any constraint assuming the shocks $\boldsymbol{\epsilon} = (\epsilon_1, .., \epsilon_n)$ to be known by the designer; it is obtained by maximizing welfare for each value of $\boldsymbol{\epsilon}$ separately. First best requires two conditions. First, the cost of achieving the total emissions volume should be minimized, which requires that private marginal costs of abatement should be equalized across firms ex post. Second, the optimal total emissions volume should emerge, which requires that this common private marginal cost of abatement should be equalized to the social marginal benefit of abatement. These conditions are referred to as *cost efficiency* and *volume efficiency* respectively. Formally, given $\boldsymbol{\epsilon}$, the first best emission volumes $x_i(\boldsymbol{\epsilon})$, for $i \in N$, satisfy:[14]

$$\begin{aligned} b_i\left(z_i + \epsilon_i - x_i(\boldsymbol{\epsilon})\right) &= m(\boldsymbol{\epsilon}) \text{ for all } i && (5) \\ m(\boldsymbol{\epsilon}) &= \nu + AX_N(\boldsymbol{\epsilon}) && (6) \end{aligned}$$

Explicit expressions are easily obtained (the computation is detailed in Appendix A). We state them in difference with respect to the allocation obtained when all shocks are equal to their mean : $\epsilon_i = 0$ for all i. Let us denote $x_i^* = x_i(\mathbf{0})$, for $i \in N$ and $m = m(\mathbf{0})$. We obtain:

$$\begin{aligned} x_i^* &= z_i - \frac{m}{b_i} \text{ for all } i \in N \text{ with } m = \frac{b_N(\nu + Az_N)}{A + b_N} && (7) \\ x_i(\boldsymbol{\epsilon}) &= x_i^* + \epsilon_i - \frac{Ab_N}{b_i(A + b_N)}\epsilon_N \text{ for all } i \in N \text{ and } m(\boldsymbol{\epsilon}) = m + \frac{Ab_N}{A + b_N}\epsilon_N. && (8) \end{aligned}$$

These expressions are linear in the shocks and the flexibility parameters. So using the notation $X_S = \sum_{i \in S} x_i$, we derive from (8):

$$X_S(\boldsymbol{\epsilon}) = X_S^* + \epsilon_S - \frac{Ab_N}{b_S(A + b_N)}\epsilon_N. \qquad (9)$$

In particular, the first best optimal aggregate volume of emissions is given by: $X_N(\boldsymbol{\epsilon}) = X_N^* + \frac{b_N}{A+b_N}\epsilon_N$. It is random and follows the aggregate shock on total gross emissions ϵ_N with a dampening coefficient $\frac{b_N}{A+b_N}$ smaller than 1. This coefficient reflects the strength

[14] We assume that all the z_i are large enough compared to $\frac{\nu}{b_i}$ so that the first best allocation as well as those considered later on produce positive values for the net emissions. This avoids uninteresting corner solutions.

of decreasing returns in the aggregate abatement technology (the slope of the aggregate marginal abatement cost curve b_N) and the slope of the marginal abatement benefit curve (A). When b_N is small relative to A, variations in the level of net emissions induce small changes in the aggregate marginal abatement cost but large swings in the social marginal abatement benefit: therefore, preserving the equality between marginal abatement cost and benefit in the presence of shocks requires to absorb most of these shocks through abatement so that net emissions do not fluctuate much. Hence, the dampening coefficient is small. The stronger decreasing returns in abatement, the steeper marginal abatement costs and so, the larger the proportion of aggregate shocks on gross emissions that is passed on into net emissions so that as to maintain efficiency.

As is well known, when there is no uncertainty, the first best allocation (x_i^*) can be reached easily either by imposing a tax on all firms ($T = N$) or by organizing an ETS ($Q = N$) provided that the tax level or the level of quotas are well chosen so as to induce the optimal common marginal abatement cost m. More generally, the optimal allocation can be reached through any scope, since the policy can be chosen to induce the marginal cost m. This neutrality result no longer holds in the presence of uncertainty: the choice of the scope matters.

2.2 Optimal tax and quota levels given a scope

Before determining the optimal scope of intervention, we analyze the optimal policy for any given scope (T, Q). Given a scope (T, Q), the optimal policy consists in the optimal tax rate to be imposed on non-ETS firms and the optimal quota allocated to the ETS, anticipating that the realized shocks will determine the emissions volume of the taxed sector and the transactions and prices within the ETS.

For a policy $(t, \overline{X}_Q)$, the amount emitted by a firm is given by $b_i(z_i + \epsilon_i - x_i) = \tau$, with $\tau = t$ in the taxed sector and $\tau = p$ in the ETS. Aggregating over firms, the amount emitted by the taxed sector and the price on the ETS, both random, are given by:

$$\widetilde{X}_T = z_T - \frac{t}{b_T} + \widetilde{\epsilon}_T \text{ and } \widetilde{p} = b_Q \left(z_Q + \widetilde{\epsilon}_Q - \overline{X}_Q \right). \tag{10}$$

The marginal abatement cost is t for the firms in T and $\widetilde{p}$ for firms in Q; so, expected marginal abatement costs are :

$$t \text{ for } i \in T \text{ and } b_Q(z_Q - \overline{X}_Q) \text{ for } i \in Q. \tag{11}$$

The first proposition characterizes the optimal policy $(t, \overline{X}_Q)$ for a given scope.

Proposition 1. *Given a scope (T, Q), the optimal level of the tax on T is equal to m and the optimal quota on Q is $\overline{X}_Q = X_Q^*$. For this optimal policy, the expectation of the emissions volume by firms in T is equal to X_T^*, and the expectation of the price level on the ETS is m.*

Proof. The proof is given in Appendix A. □

Proposition 1 shows that the optimal policy is set so as to equalize all firms' expected marginal abatement costs to m, the social marginal benefit of abatement at the optimal level absent uncertainty, and it yields in expected terms the first best optimal emission levels absent uncertainty.

More precisely, the optimal tax is set equal to m and the quota on the ETS is set equal to the aggregate optimal emissions volume of the firms under the cap-and-trade absent uncertainty. At the optimal policy, the amount emitted by firms in the taxed sector can be written, using (10):

$$\widetilde{X}_T = X_T^* + \widetilde{\epsilon}_T \tag{12}$$

that is, the sum of all shocks on gross emissions in the taxed sector are passed on, without any dampening, into fluctuations in net emissions. By contrast, all shocks affecting the gross emissions in the ETS sector are completely wiped out by construction, but the price on the ETS is random and given by:

$$\widetilde{p} = m + b_Q \widetilde{\epsilon}_Q. \tag{13}$$

The marginal abatement cost of the firms under the tax system is m whereas it is equalized to the price on the ETS market in the ETS sector; so, they are equalized in expected terms across the two sectors.

2.3 Optimal scope

The characterization given by (5)-(6) of the first best allocation incorporates two conditions: cost efficiency and volume efficiency. Under uncertainty, whatever scope (T, Q), there is little chance that both conditions are met for all realized shocks. Cost efficiency is not satisfied in a mixed system, as marginal costs are equalized either to the tax or to the (random) competitive market price on the ETS which typically differs from the tax (when $\widetilde{\epsilon}_Q$ is not equal to 0). Instead a uniform system is cost efficient. However a uniform system is not volume efficient in general. Recall that the ex post optimal volume of emissions is $X_N^* + \frac{b_N}{A+b_N}\widetilde{\epsilon}_N$. With a uniform ETS scope, the quantity is fixed and with a tax system, the aggregate volume is $X_N^* + \widetilde{\epsilon}_N$, which reflects one-for-one the aggregate shock on gross emissions and is therefore too sensitive to it. A mixed system, on the other hand, generates an aggregate (random) volume of emissions equal to: $X_N^* + \widetilde{\epsilon}_T$, which might better replicate the variations of the ex post optimal volume.[15]

To analyze further the strength of these two effects, we provide a decomposition of the loss in welfare due to the scope relative to the ex-post optimal allocation. Given a realization $\boldsymbol{\epsilon}$ of the shocks, let $W^{T,Q}(\boldsymbol{\epsilon})$ denote the welfare associated with a given scope

[15]For a common shock θ on marginal abatement costs, i.e. if $b_i\epsilon_i = b_S\epsilon_S = \theta$ for each i and S, the optimal quantity is emitted whatever θ if $b_T = A + b_N$. This is the insight of Mandell (2008).

(T,Q), and its associated optimal policy as given by Proposition 1, and let $W^{fb}(\boldsymbol{\epsilon})$ denote the optimal welfare level for the same realization of shocks, which is associated with the ex post optimal allocation. As proved in Appendix A, the welfare loss can be written as:

$$W^{fb}(\boldsymbol{\epsilon}) - W^{T,Q}(\boldsymbol{\epsilon}) = \frac{1}{2}(A+b_N)\left(\frac{b_N}{A+b_N}\epsilon_N - \epsilon_T\right)^2 + \frac{1}{2}b_N\frac{b_Q}{b_T}\epsilon_Q^2. \tag{14}$$

The first term corresponds to the loss due to a sub-optimal aggregate emissions volume, as $\frac{b_N}{A+b_N}\epsilon_N - \epsilon_T$ corresponds to the difference between the optimal volume and the volume emitted under the scope (T,Q). The second term corresponds to the loss due to the inefficient allocation of this total emissions volume across firms due to the differences in the marginal abatement cost across the taxed and the ETS sectors. The abatement of firms in T is constant and that of firms in Q is ϵ_Q. To minimize the cost of abating ϵ_Q, one should allocate it in proportion of the groups' flexibility levels, i.e. T should abate $\frac{b_N}{b_T}\epsilon_Q$ and Q should only abate $\frac{b_N}{b_Q}\epsilon_Q$. This explains why the loss is increasing in the ratio $\frac{b_Q}{b_T}$ and in the magnitude of the shocks.

Taking expectations over the distribution of the shocks, the overall expected loss can be written as:

$$W^{fb} - W^{T,Q} = \frac{1}{2}(A+b_N)\mathbb{V}\left[\frac{b_N}{A+b_N}\widetilde{\epsilon}_N - \widetilde{\epsilon}_T\right] + \frac{1}{2}b_N\frac{b_Q}{b_T}\mathbb{V}[\widetilde{\epsilon}_Q]. \tag{15}$$

This expression depends on the parameters determining the reactions of the firms and the consumers welfare, i.e. the slopes of the marginal abatement cost curves and of the marginal abatement benefit curve, and the shocks. The term affecting the expected marginal abatement costs and the expected marginal abatement benefit do not appear since we consider the optimal policy.

Our objective is to understand the factors that favor a uniform system and those that favor a mixed system.[16] For that purpose, let us first determine the best uniform system. Applying expression (15) to uniform systems, the loss due to a misallocation of emissions across the sectors (the second term) is null (since $\mathbb{V}[\epsilon_Q] = 0$ if $T = N$ and $\frac{1}{b_T} = 0$ by convention if $Q = N$) and the loss corresponding to the aggregate volume is, up to the factor $\frac{1}{2(A+b_N)}\mathbb{V}[\widetilde{\epsilon}_N]$, equal to A^2 for a system with a uniform tax ($T = N$) and to b_N^2 for a uniform ETS system ($T = \emptyset$). Thus only the value of A relative to b_N matters and the best uniform system is determined, as in the case of a single firm treated in Weitzman (1974): if the slope of the marginal abatement benefit curve is steeper than that of the aggregate marginal abatement cost, making a mistake in the level of emissions is socially more

[16] The exact determination of the optimal scope is difficult, except in specific cases. The choice variable, the scope, is a binary partition of N, which makes the optimization problem one over discrete variables, thereby preventing the use of differential techniques. For example, when all firms are symmetric, what matters is the number of firms under each regime, and the optimization problem is one over integers. Another issue is that the objective function might not to be well-behaved, even when relaxing the integer constraint.

costly than not minimizing the cost of abating, so that a uniform cap-and-trade system dominates, and conversely.

Now, comparing a mixed system with the best uniform system yields the following proposition.

Proposition 2. *Denoting* $\mathcal{A} \equiv \min\{A, b_N\}$*, the difference between the welfare associated to the best uniform system,* W^{unif}*, and that associated to the scope* (T, Q)*,* $W^{T,Q}$*, is given by* $\frac{1}{2}F(T, Q)$ *where*

$$F(T, Q) = \left\{(A - \mathcal{A})\mathbb{V}[\widetilde{\epsilon_T}] - 2\mathcal{A}\ cov(\widetilde{\epsilon_T}, \widetilde{\epsilon_Q}) + (b_N - \mathcal{A})\mathbb{V}[\widetilde{\epsilon_Q}]\right\} + b_N \frac{b_Q}{b_T}\mathbb{V}[\epsilon_Q]. \tag{16}$$

Thus, the mixed system (T, Q) *is better than any uniform system if and only if* $F[T, Q] < 0$*.*

The term within the large brackets in (16) corresponds to the difference in the fluctuations around the first best volume between the best uniform system and (T, Q); the second term corresponds to the cost inefficiencies.

There is a fundamental asymmetry between T and Q in the criterion incorporated in (16), which is due to the external effects imposed by a firm on others in the ETS. Adding up a firm i in the ETS increases the flexibility of the ETS sector and therefore its ability to absorb a given shock with less fluctuations in the ETS price. Moreover, if this additional firm in the ETS suffers no shock, whether it is assigned to the sector under the tax or to the ETS has no impact on the fluctuations in total emissions and overall, adding up firm i in the ETS improves cost efficiency without affecting volume efficiency. Formally, this can be checked by computing the difference in our criterion when firm i, such that $\widetilde{\epsilon_i} \equiv 0$ is shifted from T to Q:

$$F(T, Q) - F(T - i, Q + i) = b_N\left(\frac{b_Q}{b_{N-Q}} - \frac{b_{Q+i}}{b_{N-Q-i}}\right)\mathbb{V}[\widetilde{\epsilon}_Q].$$

The difference is positive as $b_Q > b_{Q+i}$ and $b_{N-Q} < b_{N-Q-i}$, hence it is always worth shifting i to Q. This argument indicates that for the optimal scope, the ETS sector should incorporate all firms that are hit by small shocks.

The next corollary provides a necessary condition for a mixed system to be optimal.

Corollary 1. *A mixed system* (T, Q) *is welfare improving over the best uniform system only if the following holds:*

$$b_N \frac{b_Q}{b_T}\mathbb{V}[\widetilde{\epsilon}_Q] < 2\mathcal{A}\ cov(\widetilde{\epsilon}_T, \widetilde{\epsilon}_Q). \tag{17}$$

In particular, the co-variance cov $(\widetilde{\epsilon}_T, \widetilde{\epsilon}_Q)$ *must be positive.*

Proof. Obvious since the terms $(A - \mathcal{A})\mathbb{V}[\widetilde{\epsilon_T}]$ and $(b_N - \mathcal{A})\mathbb{V}[\widetilde{\epsilon_Q}]$ in F are non-negative. □

As we have seen, the loss due to the cost inefficiency of a mixed system is increasing in the ratio $\frac{b_Q}{b_T}$ and in the magnitude of the shocks, $\mathbb{V}[\widetilde{\epsilon}_Q]$. Condition (17) is a necessary condition for cost inefficiency of a mixed system to be smaller than the gain in volume efficiency generated by the fact that the mixed system tracks the fluctuations of the first best allocation better than the best uniform system. This is possible only if ϵ_T and $\frac{b_N}{b_N+A}\epsilon_N$ are highly correlated that is, if the co-variance of the shocks between the two sectors is positive and strong enough.

As a result of Corollary 1, the optimal scope is the best uniform system when the shocks affecting the firms are two-by-two negatively correlated or independent. The case of two-by-two negatively correlated or independent shocks is however a rather implausible assumption.

Corollary 2. *If a mixed system is optimal for $A \neq b_N$, then a mixed system is optimal in the critical case where $A = b_N$. In that critical case, a mixed system (T, Q) is welfare improving over the best uniform system if and only if:*

$$b_Q \mathbb{V}[\widetilde{\epsilon}_Q] < 2b_T \; cov(\widetilde{\epsilon}_T, \widetilde{\epsilon}_Q), \tag{18}$$

Proof. From Corollary 1, the co-variance $\text{cov}(\widetilde{\epsilon}_T, \widetilde{\epsilon}_Q)$ must be is positive for the mixed scope (T, Q) to be optimal. When this condition is satisfied, the value of $F[T, Q]$ decreases in A as long as $A < b_N$, and increases for $A > b_N$. The first statement follows. As for the second, it is immediate since in the critical case $F(T, Q) = -2b_N \text{ cov}(\widetilde{\epsilon}_T, \widetilde{\epsilon}_Q) + b_N \frac{b_Q}{b_T} \mathbb{V}[\widetilde{\epsilon}_Q]$. □

It is precisely when the two uniform systems are equivalent that they are most likely to be dominated by a mixed system: to check that a mixed system is optimal in this critical case, it is sufficient to exhibit a scope, possibly different from the optimal one, for which condition (18) is satisfied. When A differs from b_N, condition (18) is only necessary. There is, however, little evidence on how A and b_N differ worldwide as the aggregate marginal abatement benefits are far from being consensual even within the scientific community. So, we view the critical case as a relevant benchmark case.

Condition (18) is rather mild, as suggested by the two following situations. In the first situation, firm i_0 is affected by a low variance shock that is highly correlated with the high-variance shocks affecting (some) other firms in the economy. Then, condition (18) holds for $Q = \{i_0\}$ provided that firm i_0's flexibility is not too small compared to the other firms'. In the second situation, each activity $i \in I$ in the economy is duplicated into two sub-activities, $(i, 1)$ and $(i, 2)$, with $b_{(i,1)} = b_{(i,2)}$ and $\widetilde{\epsilon}_{(i,1)} = \widetilde{\epsilon}_{(i,2)} = \widetilde{\epsilon}_i$. Then, the scope (T, Q), with $T = \{(i, 1), i \in I\}$ and $Q = \{(i, 2), i \in I\}$, satisfies (18). This example contains an important insight if we consider that the economy consists in several sectors populated by similar firms, active on the same markets and subject to the same shocks. Absent any technical or political constraints, it is beneficial in such a situation to split each sector so that a fraction of firms is under the cap-and-trade mechanism and the rest is subject to the tax: then, ϵ_T corresponds to a scaled-down version of the aggregate shock ϵ_N, where the scale can be adjusted so as to equal $\frac{b_N}{A+b_N}$ (1/2 in the critical case). Volume efficiency is maximized and dominates cost inefficiency. Such a mixed system would, however, have to treat similar firms differently, which raises obvious issues about acceptability.

We now illustrate Proposition 2 with two examples in which the optimal scope can be computed.

Example with symmetric firms. All firms exhibit the same sensitivity to shocks, $b_i = b$ for all i, and the shocks ϵ_i have identical variance σ^2 and are two-by-two correlated with an identical correlation coefficient equal to ρ. The symmetry assumption implies $b_S = \frac{b}{|S|}$. Furthermore the value of $F[T, Q]$ only depends on the number of firms in T and Q.

Let us determine the conditions under which a mix system is optimal. From Corollary 1, the correlation ρ has to be positive. Let us start with the critical case in which $A = b_N = b/n$. Simple computation yields that function F is equal to :

$$\frac{\sigma^2}{n}(n-q)\left[-\rho q + 1 - \rho\right],$$

where q is the number of firms in Q. Relax the integer constraint on q, F is convex in q for $\rho > 0$ and reaches its minimum at $q^* = \frac{n-1}{2} + \frac{1}{2\rho}$. This is always positive, and larger than 1 if $n \geq 3$. It is thus always optimal to have an ETS. There is also a taxed sector if $q^* < n$, or equivalently if $\rho \geq \frac{1}{n-1}$. Thus a mixed system is likely to be optimal.

When we are not in the critical case, there is a force towards extending the ETS or the taxed sector depending on whether A is larger or smaller than b/n.

Example with a common shock on marginal abatement costs. Let us consider the polar situation in which the same shock affects all firms' marginal abatement cost curves, that is: $b_i\widetilde{\epsilon}_i = \widetilde{\theta}$ for all i. The following corollary characterizes the values for which a mixed system is optimal.

Corollary 3. *Assume a common shock on marginal abatement costs, $b_i\widetilde{\epsilon}_i = \widetilde{\theta}$ for all i. An optimal scope is a mixed system for $b_N < 2A < 2\max_{i\in N} b_i$. Up to indivisibilities, an optimal scope satisfies*[17] *$2A = b_T$ when $2A < \max_{i\in N} b_i$ and is reduced to a single firm (with maximum b_i) in the taxed sector for $A < \max_{i\in N} b_i < 2A$. Otherwise, the optimal scope is a uniform tax system for $b_N > 2A$ and a uniform cap-and-trade scheme for $\max_{i\in N} b_i < A$.*

[17] Any T such that $2A = b_T$ is optimal and, if there is no such T, b_T at the optimal scope is either the greatest value smaller than $2A$ or the smallest one greater than $2A$.

Proof. When $b_i\widetilde{\epsilon}_i = \widetilde{\theta}$ for all i, $b_S\widetilde{\epsilon}_S = \widetilde{\theta}$ for any subset $S \subset N$. Thus, with $\sigma^2 = \mathbb{V}[\epsilon]$, (16) can be written as:

$$\begin{aligned} F(T,Q) &= \sigma^2\{\frac{A}{b_T^2} + \frac{1}{b_Q} - \mathcal{A}[\frac{1}{b_N^2}]\} \\ &= \sigma^2\{\frac{1}{b_T}(\frac{A}{b_T} - 1) + \frac{1}{b_N}(1 - \frac{\mathcal{A}}{b_N})\} \end{aligned}$$

From this expression, an optimal mixed system must minimize $\frac{1}{b_T}(\frac{A}{b_T} - 1)$, which, up to indivisibilities, yields $b_T = 2A$ and the value $\sigma^2\{\frac{1}{4A} + \frac{1}{b_N}(1 - \frac{\mathcal{A}}{b_N})\}$. This expression must be negative for the mixed system to dominate the uniform systems. The proposition follows. □

In the critical case, where $b_N = A$, a mixed system is surely optimal (as $b_N \leq b_i$ for each i). It is reasonable to assume that there are firms with very small flexibility parameters, i.e. large b_i, so that $\max_{i \in N} b_i$ is very large. Assuming this is the case, a mixed system is optimal whenever $b_N < 2A$, and otherwise it is a uniform tax system. In this polar case of a common shock on marginal abatement costs, only the aggregate flexibility of each sector matters, and there is no general result about the values of b_i that should be included in the set of firms under the tax or in the ETS, provided the optimal aggregate flexibility in each sector is reached.

2.4 Discussion and extensions

Let us first discuss how our normative results in the previous sub-section can be used to assess the performances of actual climate policies. The necessary condition for a mixed system to be better than the best uniform system is that it induces random net emissions (from the taxed sector) that are positively correlated with the variations of its ETS price. From (12), the fluctuations in the net emissions are equal to the shocks affecting the gross emissions of the firms that are subject to the tax, while from (13) the price fluctuations on the ETS market are positively correlated with the shocks affecting the gross emissions of the ETS firms. Analysts of existing emissions trading systems often consider that the price fluctuations of allowances are mainly driven by two factors: macroeconomic shocks and weather shocks. Macroeconomic shocks affect basically all sectors that emit GHG in a strongly correlated way. So, it is probably a good signal on the design of the European mixed system that the price of allowances is highly correlated with macroeconomic conditions, and therefore with shocks affecting gross emissions in non-ETS sectors (See Chevalier (2011)). By contrast, climatic shocks affect mostly the power and heat generation industries, as well as agriculture and farming in a milder way; they may however have only very limited impact on transportation and energy intensive industries. This suggests that a mixed system that would impose a cap-and-trade regulation only on the power and heat generation industries would probably be suboptimal since the co-variance term in (18) would be rather low, in particular if agriculture and farming were included in the ETS or if they represented a limited fraction of the economy. In other terms, in an area where agriculture and farming are important and are affected by strong weather fluctuations, e.g. perhaps as in Australia or North America, it may be preferable not to include them in a ETS mostly targeting power plants.

Targeting a volume of emissions Our results can easily be amended to a setting in which the designer is subject to an additional constraint on its expected emissions. Indeed, areas that, as the European Union, have signed the Kyoto protocol have to meet an additional emissions target, which constrains their choice of an optimal scope and associated policies. Viewing such an emissions target as a strictly-enforced cap on (expected) emissions amounts to introducing the constraint: $\mathbb{E}[\widetilde{X}_N] \leq \overline{X}$. Of course, if the expected volume that the country chooses without constraint is less than the target, $\overline{X} \geq X_N^*$, nothing is changed. When $\overline{X} < X_N^*$, we can show that this additional constraint only modifies the deterministic parts of the net emissions and of the policy levels; however, the effects of the shocks on welfare are unchanged, so that the determination of the optimal scope is the same as before.[18]

More precisely, consider first the optimal policy given a scope. The designer's problem then becomes one of minimizing the cost of achieving the given target $\overline{X}$, and the solution simply amounts to equalizing marginal abatement costs across the taxed sector and the ETS sector.[19] Formally, the optimal policy consists in fixing $(t, \overline{X}_Q)$ such that: $t = b_Q(z_Q - \overline{X}_Q) = b_N(z_N - \overline{X})$. It then follows that the ETS price is given by $\widetilde{p} = t + b_Q\widetilde{\epsilon}_Q$ and aggregate emissions are given by: $\widetilde{X}_N = \overline{X} + \widetilde{\epsilon}_T$.

Consider now the choice of the scope. Setting the policy as just described, the deterministic parts of net emissions and of the ETS price are identical whatever the scope. So, the levels of expected welfare across various scopes only differ through the impact of the shocks and the function F is still the criterion to compare the scopes. It follows that the determination of the optimal scope does not depend upon whether we take into account an emissions target constraint or not: it only depends upon the fact that expected marginal abatement costs are equalized across all firms.

Accounting for the impact of the climate policy on the goods markets. So far we have assumed no interaction between the goods markets and the environmental policy. In particular, we have assumed that gross emission levels are not affected by the fiscal instruments. Even in the absence of shocks, this assumption is debatable. Though the interaction surely exists, there is no consensus about its magnitude, as testified by the vivid debate on the impact of the environmental tax policies on growth (see Stern,

[18]From an analytic point of view, adding the constraint to the program modifies the objective by adding a linear term of the form $\lambda(\overline{X} - \mathbb{E}[\widetilde{X}_N])$ for an appropriate multiplier λ. The result follows as linear terms have no impact in our analysis. Formally, this simply modifies the value of m by substituting $\nu + \lambda$ to ν.

[19]This is the view adopted in e.g. Heindl - Wood - Jotzo (2014) or Wood *et al.* (2013).

2007, Nordhaus, 2007) or, in a shorter term perspective, on the labor market - the double dividend debate (Bovenberg and Goulder, 1996). At a micro level, Martin, de Preux and Wagner (2014) find that a carbon tax has an impact on energy use, but do not find a significant effect on employment and production.

Also, we have assumed the shocks on emission levels to be independent of those affecting the goods markets. While this assumption is appropriate for modeling shocks on the abatement technologies, it does not fit when shocks bear e.g. on the price of energy. In that case, the gross emission level is related to the quantity of energy input. We indicate here how the analysis carries over more generally provided some approximations are valid.

First, assume firms recognize the impact of fiscal instruments when contemplating their production decisions. Firms subject to the ETS make their production decisions on the basis of the expected ETS price, i.e., before they observe the realization of the ETS price, but of course they make their abatement and purchase of permits decisions after observing the ETS price.

Second, assume the impact of the fiscal instruments on the *variations* in gross emissions around their mean are negligible. Specifically, given a tax level or an expected permit price τ, let $z_i(\tau)$ be i's gross emission level under certainty at the equilibrium of the market, i.e. accounting for the impact of τ on the goods market equilibrium. Shocks on production lead firms to adjust their production level resulting in a gross emission level, say $z_i(\tau, \eta)$ if η denotes the production shock. We assume that we can neglect the impact of τ on the variation of gross emissions around the value without shock, i.e. $z(\tau, \eta) - z(\tau)$ is a random variable independent of τ, which has a null expectation. Thus, $z(\tau, \eta)$ can be written as $z(\tau) + \epsilon$.

Third, welfare is quadratic in the choice variables. This assumption was already made on the emission costs and damage; hence it is extended to the surplus on the goods markets.

Under these three assumptions,[20] given a scope, the optimal policy (tax and quota) is the one that maximizes the welfare in the absence of shocks, as in Proposition 1. This implies that the tax is independent of the scope and the quota is set so that the expected permit price is equal to that optimal tax. The only difference with the previous analysis here is that this optimal policy accounts for the equilibrium effect on the goods markets. The analysis of the scope (see the next subsections) then carries through: since, whatever the scope, the tax and expected emission price are set to the optimal tax without uncertainty, the gross emission levels are independent of the scope.

[20] The key point in the argument is that marginal welfare with respect to a marginal variation in fiscal instruments is linear in the shocks. Its expectation is independent of the shocks and equal to the level in the absence of shocks (the certain equivalent).

2.5 The firms' viewpoint

We consider here the point of view of the firms on the design of the scope. More precisely, we ask the following question: given a scope, would a taxed firm rather be subject to cap-and trade, and conversely ? For that we compare firms' expected profits under both systems.

Given the realized ϵ_i simple computation yields that i's profit when it faces the cost τ (dropping the independent term ξ_i) is: $\Pi_i = \frac{\tau}{2b_i}(\tau - 2b_i(z_i + \epsilon_i))$. Taking expectations over the shocks, we obtain the expected profits at the optimal policy. For a firm i subject to the tax, $\widetilde{\tau} = t = m$ and its expected profit is:

$$\mathbb{E}\left[\widetilde{\Pi}_i\right] = \frac{m}{2b_i}(m - 2b_i z_i). \tag{19}$$

For a firm i subject to cap-and-trade, the cost of emission is equal to $\widetilde{\tau} = \widetilde{p} = m + b_Q\widetilde{\epsilon}_Q$, which gives the expected profit:

$$\mathbb{E}\left[\widetilde{\Pi}_i\right] = \frac{m}{2b_i}(m - 2b_i z_i) + \frac{1}{2b_i}\,\mathrm{cov}(b_Q\widetilde{\epsilon}_Q, b_Q\widetilde{\epsilon}_Q - 2b_i\widetilde{\epsilon}_i). \tag{20}$$

The profit for a firm under tax is independent of the scope, as there is no external effect across firms and furthermore the tax level stays equal to m whatever the scope. Instead, the profit for a firm under cap-and-trade depends on the scope as the price depends on the firms under the ETS.

Consider a firm i under the ETS. We assume that i compares its current profit to the profit it would achieve under the tax.[21] For i in Q, i's profit is larger under the cap-and-trade than under the tax if and only if the co-variance term in (20) is positive. The next proposition suggests that the most likely situation seems to be that firms prefer to be subject to a tax than to be included in the cap-and-trade mechanism.

Proposition 3. *Whatever the shocks structure, at least one firm under the cap-and-trade mechanism would prefer to be subject to the tax, and this is the case for all firms when (a) the shock is common to all firms' marginal abatement costs, or (b) firms are symmetric: the shocks have identical variance and correlation and all flexibility parameters are identical ($b_i = b$ for all i).*

Proof. The proof is given in Appendix A. □

From a political economy perspective, we can therefore expect that firms will likely be opposed to the design of a cap-and-trade mechanism and that those subject to it will actively lobby so that the system be abandoned or they be taken out of it.

[21] For a firm under tax, i in T, the reverse criteria says that i compares its current profit to the profit it would achieve if it joined the ETS, when $Q \cup \{i\}$ is the ETS sector. This means that the firm accounts for its impact on the price of the ETS. An alternative assumption would be that the firm computes its profit under the observed price, when Q is the ETS sector. The difference is likely to be negligible for most firms but not for a large electricity firm like EDF or for an industry.

Finally the comparison of the firms' profits in the two sectors holds more generally in the case of an emissions target. More precisely, given that expected marginal abatement costs are equalized across the two sectors, i.e. $b_Q(z_Q - \overline{X}_Q) = t$, the firms' profits are given by the same type of expressions as in (19) and (20), where the constant term is modified but still identical across sectors, so that the comparison of the profits only relies on the variance terms.

3 Climate policies in a world with several areas

A prominent issue in the design of a worldwide regulation of GHG emissions is that various countries around the world, or more generally various areas, have chosen their modes of regulation separately. Inefficiencies naturally arise from the non-cooperative design of climate *policies* in the absence of uncertainty. We first show in this section how they are reinforced by the introduction of uncertainty. We then discuss how the decentralized design of *scopes* matters: it does not impact expected emissions but it affects welfare through the shocks and we analyze the interaction among areas at this scope design stage. Finally, the analysis enables us to provide a simple and quantitative assessment of a proposal to merge the various ETS across areas : such an agreement constitutes a clear improvement over the non-cooperative benchmark and it seems easily implementable.

Consider a world consisting of several areas indexed by $\alpha \in \mathcal{W}$. Consumers are concerned with the worldwide emissions volume. Consumers' welfare in area α is:

$$S^\alpha(X) = \lambda^\alpha - \nu^\alpha X - \frac{A^\alpha}{2}X^2,$$

where $X = \sum_{\alpha\in\mathcal{W}} X^\alpha$ is the worldwide level of emissions. So, consumers' surplus worldwide is given as before in (3) with $\nu = \sum_{\alpha\in\mathcal{W}} \nu^\alpha$, $\lambda = \sum_{\alpha\in\mathcal{W}} \lambda^\alpha$ and $A = \sum_{\alpha\in\mathcal{W}} A^\alpha$. As in the previous section, x_i^* for $i \in \cup_{\alpha\in\mathcal{W}} N^\alpha$ denotes the (worldwide) optimal net emissions volume absent uncertainty and m denotes the firms' common marginal abatement cost which is equal to the worldwide social abatement benefit.

As in the previous section, the first best is not implementable because the regulation is designed before the realization of the shocks. When the scope in each area is given, (T^α, Q^α) for area $\alpha \in \mathcal{W}$, and climate policies are chosen by the world planner, a simple adaptation of Proposition 1 shows that these policies consist in fixing identical tax rates equal to the common marginal abatement cost at the worldwide optimum absent uncertainty, $t^\alpha = m$ for any area $\alpha \in \mathcal{W}$, and quotas equal to the optimal emissions volumes absent uncertainty, $\overline{X}^\alpha = X^*_{Q^\alpha}$. Alternatively, these climate policy choices would be the ones made under a cooperative scenario across areas, for the given scopes. The following section instead considers a non-cooperative scenario.

3.1 Decentralized choice of climate policies

Suppose that, in each area $\alpha \in \mathcal{W}$, the scope (T^α, Q^α) is given and the local policy designer is concerned with the expected welfare of the entities in the area only, and each local policy designer chooses its policy in a non-cooperative way.

Given scope (T^α, Q^α) in area $\alpha \in \mathcal{W}$, let t^α denote the tax rate, $\overline{X^\alpha}$ the quota on the local ETS, $\widetilde{\Pi}_i$ firm i's profit and $\widetilde{p}^\alpha$ the ETS price in area α. The expected welfare in area α is composed of the surplus of its consumers, the profit of the firms in the area and the public revenues from the taxes and the sales of permits:

$$\mathbb{E}\left[S^\alpha(\widetilde{X}^\alpha + \widetilde{X}^{-\alpha}) + \sum_{i \in N^\alpha} \widetilde{\Pi}_i + \widetilde{p}^\alpha \overline{X^\alpha} + t^\alpha \sum_{i \in T^\alpha} \widetilde{x}_i\right].$$

where $\widetilde{X}^{-\alpha}$ stands for the sum of net emissions in all areas except α. Let $m^\alpha = \nu^\alpha + A^\alpha X^*$ denote the marginal abatement benefit in area α at the global optimum without uncertainty, and note that: $m = \sum_{\alpha \in \mathcal{W}} m^\alpha$. The following proposition characterizes the policy choices in the Nash equilibrium of the game in which all areas choose independently and simultaneously their climate policies, given the existing scopes.

Proposition 4. *Given scopes (T^α, Q^α) for $\alpha \in \mathcal{W}$, the climate policies at the non-cooperative equilibrium are given by:*

$$t^\alpha = \nu^\alpha + A^\alpha \mathbb{E}[X], \; \overline{X^\alpha} = z_{Q^\alpha} - \frac{t^\alpha}{b_{Q^\alpha}} \tag{21}$$

$$\textit{and } \mathbb{E}[X] = X^* + \frac{\sum_{(\beta,\gamma) \in \mathcal{W}^2, \beta \neq \gamma} \left[\frac{m^\beta}{b_{N^\gamma}}\right]}{1 + \sum_{\beta \in \mathcal{W}} \left[\frac{A^\beta}{b_{N^\beta}}\right]}. \tag{22}$$

This results in an excess in the expected worldwide emissions volume compared to the worldwide social optimum, i.e. $\mathbb{E}[X] > X^*$, and $t^\alpha > m^\alpha$.

Proof. See Appendix B. □

Not surprisingly, in equilibrium, areas choose policies that lead ultimately to an excess of expected emissions from a worldwide perspective.

According to the first equation in (21), the marginal abatement costs in the taxed sector are equalized to the expected marginal abatement benefit in area α and, according to the second one, the expected price in the ETS is equalized to the tax rate ; so expected marginal abatement costs are equalized across firms within each area. From (22), the expected aggregate emissions volume is independent of uncertainty and of the scopes fixed in each area. As a result, the tax rates are independent of the scopes, hence the expected marginal abatement costs and the expected emissions volume within each area as well.

It follows from this discussion that the policy choices produce the same outcome in expected terms as the ones that would be chosen at a Nash equilibrium in the absence of uncertainty. Inefficiencies associated to this Nash equilibrium come from two channels : first, there is no reason for the equalization of the marginal abatement costs or benefits across areas, as the levels t^α are likely to differ across areas; and second, there is an excess of emissions compared to the worldwide optimum (even if areas are all symmetric). This excess comes from the fact that each area is concerned with the surplus of its own consumers only, as reflected by the expression for the tax levels in (21).

Though there is an overall excess in emissions at the Nash equilibrium, it might not be the case for each area. It is possible that one area chooses a more stringent climate policy than what is worldwide optimal to 'compensate' for the lax policy of other areas. More precisely, Appendix B proves that in equilibrium and with two areas α and β, it is possible that $t^\alpha = b_{Q^\alpha}(z_{Q^\alpha} - \overline{X^\alpha}) > m$ if A^α and ν^α are large compared to A^β and ν^β, which implies that $m^\alpha >> m^\beta$. So, when the two areas are very different, the area with the largest concern for emissions at the worldwide optimum may compensate for the lax policy of the other area. Moreover, in that case, $\overline{X^\alpha} < X^*_{Q^\alpha} \Leftrightarrow m < t^\alpha$: that is, both area α's taxed sector and cap-and-trade sector are distorted in the same direction.

In line with the discussion of sub-section 2.5, additional Kyoto-type constraints on the expected volume of emissions in each area can be introduced in the analysis with only minor changes. Cost efficiency within each area is obtained in equilibrium since when the expected emissions constraint is binding for a given area, this area is interested in achieving its given level of expected emissions at the lowest possible social cost. So, only total expected emissions may be affected by Kyoto-type constraints, and the equilibrium determination simply relies on "capping" each best-response by the allocated quota in each area. Given the nature of best-responses (decreasing with slopes smaller than 1), if each area is imposed a quota smaller than its expected volume of emissions in the unconstrained equilibrium, all areas will emit their allocated quotas in the constrained equilibrium. In particular, if it is the case that all areas emit more in the unconstrained equilibrium than in the worldwide optimum, Kyoto-type constraints at the levels of worldwide optimal emissions restore global efficiency at the world level. As argued previously, this occurs when areas are not too dissimilar. By contrast, if areas are quite different so that worldwide efficiency would require reducing the emissions of some but increasing the emissions of others, the imposition of Kyoto-type quotas may not be effective in restoring worldwide efficiency.

3.2 Decentralized choice of scopes

In the previous analysis, we have been silent about the determination of the scopes in each area. As we have just seen, scopes matter only because there are shocks; their determination does not impact the average emissions volume but it does impact the random component of the emissions, hence the expected welfare at the local level of each area. It is most likely that, in practice, the design of scopes in various areas has been, and still is, both a question of social welfare at the level of the area and a political economy issue in which local firms and local lobbies influence the determination of the scope. So we do not consider a non-cooperative approach in which local designers would play a game in scope design. However, in the rest of this subsection, we point out some implications of the interaction across scopes, in particular we display an interesting substitutability property.[22]

We first analyze the choice of a single area, taking as given the random emission volume of the rest of the world. To simplify notation and avoid indices, we do not index the area under consideration and denote by $Y + \widetilde{\eta}$ the emissions volume of the rest of the world. The variance of the shock is not null, since otherwise the previous analysis readily applies.

Adapting the analysis of the previous section, let W^{fb} denote the first best level for country α when it has full information on the shocks realized in its area and on the outside shock $\widetilde{\eta}$. In this situation, the overall expected loss can be written as:

$$W^{fb} - W^{T,Q} = \frac{1}{2}(A + b_N)\mathbb{V}\left[\frac{b_N\widetilde{\epsilon}_N - A\widetilde{\eta}}{A + b_N} - \widetilde{\epsilon}_T\right] + \frac{1}{2}b_N\frac{b_Q}{b_T}\mathbb{V}[\widetilde{\epsilon}_Q]. \tag{23}$$

The second term, which reflects cost inefficiency, is not changed. The first term, which corresponds to the loss due to a sub-optimal aggregate emissions volume, is changed because the first best emission level changes and accounts for outside emissions. It is given by $\frac{b_N\widetilde{\epsilon}_N - A\widetilde{\eta}}{A+b_N}$: the emissions volume within the country is reduced when that from outside is high, and conversely, thereby smoothing the total emissions volume.

The best uniform system for the area can be easily derived. Up to the factor $\frac{1}{2(A+b_N)}$, the expected loss levels are respectively $\mathbb{V}[A\widetilde{\epsilon}_N + A\widetilde{\eta}]$ and $\mathbb{V}[b_N\widetilde{\epsilon}_N - A\widetilde{\eta}]$ under a tax and a cap-and-trade systems.

The choice between the uniform systems relies not only on the comparison between the slopes of the local marginal abatement benefit curve and of the local marginal abatement cost curves but also on the shock in the other areas. More precisely, writing the difference in losses as $(A + b_N)[(A - b_N)\mathbb{V}[\widetilde{\epsilon}_N] + 2Acov(\widetilde{\epsilon}_N, \widetilde{\eta})]$, it also depends on the correlation between $\widetilde{\eta}$ and $\widetilde{\epsilon}_N$. A positive correlation increases the difference, hence tends to favor the ETS and a negative one the tax system. This is easy to understand: the impact of emissions of the taxed firms are reinforced in the former case and dampened in the latter.

This result provides an insight on the strategic choice of the uniform systems. Specifically, make the (heroic) assumption that each country chooses a uniform system and an optimal policy. Assume also that the correlations between the shocks of each pair of countries is non-negative.

From the viewpoint of country α, we have $\widetilde{\eta} = \sum_{\beta \in N_T} \widetilde{\epsilon}_{N^\beta}$ where N_T represents the set

[22]For the same reason as evoked in subsection 2.5, the imposition of constraints on the expected volume of emissions in each area does not impact the trade-offs involved in the design of the scope in each area.

of countries other than α choosing a uniform tax system. The incentives for α to choose a uniform cap-and-trade system,[23] hence increase with the set N_T, due to the positive correlation between shocks. If $0 < [b_{N^\alpha} - A^\alpha]\mathbb{V}[\widetilde{\epsilon}_N] < 2A^\alpha cov(\widetilde{\epsilon}_{N^\alpha}, \sum_{\beta\neq\alpha} \widetilde{\epsilon}_{N^\beta})]$, area α's welfare is larger with a uniform tax system than with a uniform cap-and-trade system when the other areas choose a uniform cap-and-trade and conversely when they all choose a uniform tax system. Viewed as the characterization of a best response scope design for area α, this means that there is some form of strategic substitutability in the design of scopes when only local welfare criteria drive the process. This conclusion suggests that one might observe a lot of variability in the systems adopted worldwide even though there are strong correlations and areas have similar preferences, provided social welfare considerations play a sufficiently important role in the determination of scopes.

Looking at mixed systems is more difficult and we simply illustrate this idea of strategic substitutability in the context of a symmetric model. More precisely, assume symmetry among marginal abatement costs worldwide, $b_i = b$ for all i, and consider the following stochastic structure: $\mathbb{V}[\epsilon_i] = \sigma^2$ for any $i \in N$, $\text{cov}(\epsilon_i, \epsilon_{i'}) = \rho^\alpha$ if $i, i' \in N^\alpha$ and $\text{cov}(\epsilon_i, \epsilon_{i'}) = r$ if $i \in N^\alpha$ and $i' \in N^\beta$ with $\alpha \neq \beta$. Given the symmetry across firms within an area, a scope is entirely characterized by the number of firms h^α included in the taxed sector in area α. The scope that maximizes area α's social welfare can be shown to correspond to the number h^α that minimizes:

$$A^\alpha[h^\alpha(1 + \rho^\alpha(h^\alpha - 1)) + 2rh^\alpha h^{-\alpha}] + b(1 + \rho^\alpha(n^\alpha - h^\alpha - 1)),$$

which is convex in h^α. The cross derivative with respect to h^α and $h^{-\alpha}$ being positive, it follows that the slope of area α's best response is negative. So, in this symmetric case, the property of strategic substitutability also holds.

This property of strategic substitutability leads to some applied consequences, provided that social welfare arguments are indeed critical in the design of scopes. The more inclusive the EU in its ETS, as for example through the extension of the ETS to airlines and air transportation, the stronger the forces pushing newly created ETS elsewhere in the world to be restrictive in their own design of ETS. One argument that has been often raised in favor of the pioneering decision of the EU ETS is that the EU provides a leading example that would trigger imitation elsewhere by other areas and eventually some convergence to a large worldwide ETS. This argument relies heavily on some assumption of complementarity among scope design decisions which does not correspond to our finding; one can therefore be skeptical about this leadership process and its virtuous global consequences .

[23]The uniform cap-and-trade system is the best choice for α if $[(A^\alpha - b_N^\alpha)\mathbb{V}[\widetilde{\epsilon}_{N^\alpha}] + 2Acov(\widetilde{\epsilon}_{N^\alpha}, \sum_{\beta\in N_T} \widetilde{\epsilon_{N^\beta}})]$ is positive.

3.3 Merging ETS

Given the inefficiencies that characterize the non-cooperative determination of regulatory scopes and climate policies, the question arises of how various areas could coordinate and cooperate, still assuming that no authority could enforce a worldwide agreement.

One route that has been discussed is to link the existing ETS of different areas, while preserving sovereignty of each area (country). We consider here the merging of ETS, keeping the other parameters of the climate policies -the taxes and scopes- unchanged in each area (we do not make specific assumptions on these parameters, in particular they do not necessarily form a Nash equilibrium). The merging of ETS results in a unique ETS for which the quota is equal to the sum of the quotas in the areas and the participants are all the firms initially subject to the ETS of their respective areas: that is, $\overline{X} = \sum_{\alpha\in\mathcal{W}} \overline{X}^{\alpha}$ and $Q = \cup_{\alpha\in\mathcal{W}} Q^{\alpha}$. We also assume that the revenues from the sale of the permits are allocated to each area proportionally to its original quota.

Under such scenario, the merging of the ETS has an effect only on the firms under ETS and the revenues from the sales of the permits. Indeed, the taxed sectors, the fiscal revenues from the emissions taxes and the consumers' welfare in each area are unaffected. This is due to the fact that the volume of emissions of the taxed firms is identical under separate or merged ETS. As a result, the total volume of emissions is identical in the two scenarios as well. Hence the consumers' expected surplus in each area are also identical as they depend only upon the aggregate emissions volume. The comparison between the two situations (separate ETS or merged ETS) thus boils down to the comparison of the sum of expected profits of the firms under ETS and the revenues from the sales of the permits across the two situations.

The questions we investigate are: (1) Does the merging of ETS induce an increase in worldwide social welfare ? (2) Does the merging of ETS induce an increase in welfare in each area ? (3) Can we evaluate the welfare changes due to merging ETS ?

The answer to the first question is positive since doing so enables one to equalize marginal costs across all firms within Q for each realization of the shocks, hence a welfare gain, while all the rest remains unchanged, in particular total emissions worldwide. The answer to the second question is less obvious, but turns out to be positive, as stated in the following proposition, which also answers the third question.

Proposition 5. *Fix the scopes (T^{α}, Q^{α}), the tax levels t^{α} and the quotas $\overline{X}^{\alpha}$ in each area $\alpha \in \mathcal{W}$. Then, merging the ETS systems strictly increases area α's social welfare for any realization of the shocks for which its separate ETS price $\widetilde{p}^{\alpha}$ and the merged ETS price $\widetilde{p}$ differ. Furthermore, the expected gain in area α increases in $\mathbb{E}[(\widetilde{p}^{\alpha} - \widetilde{p})^2]$.*

Proof. See Appendix B. □

The proposition asserts that merging ETS can only be beneficial for *each* area, even in the absence of compensatory transfers across areas. The argument is simple and we provide it below.

We show that, for each realized shocks, the sum of the expected profits of the firms under ETS and the revenues from the sales of the permits in an area increases (resp. is unchanged) provided the price on the merged ETS differs from the separate ETS price (resp. is equal). For given shocks $\boldsymbol{\epsilon}$, let $x_i^\alpha = z_i + \epsilon_i - \frac{p^\alpha}{b_i}$ and $x_i = z_i + \epsilon_i - \frac{p}{b_i}$ denote firm i's chosen emissions for a firm in area α, under separate and merged ETS respectively, and let $\Pi_i^\alpha(\boldsymbol{\epsilon}) = \xi_i - \frac{b_i}{2}(z_i + \epsilon_i - x_i^\alpha)^2 - p^\alpha x_i^\alpha$ and $\Pi_i(\boldsymbol{\epsilon}) = \xi_i - \frac{b_i}{2}(z_i + \epsilon_i - x_i)^2 - p x_i$ the corresponding profits attained in each situation. Since x_i maximizes firm i's profit when facing price p, the following must be true:

$$\Pi_i(\boldsymbol{\epsilon}) \geq \xi_i - \frac{b_i}{2}(z_i + \epsilon_i - x_i^\alpha)^2 - p x_i^\alpha = \Pi_i^\alpha(\boldsymbol{\epsilon}) + (p^\alpha - p)x_i^\alpha.$$

Summing over $i \in Q^\alpha$ and assuming all permits are sold under the initial situation, i.e. assuming that $\sum_{i \in Q^\alpha} x_i^\alpha = \overline{X}_{Q^\alpha}$, it follows:[24]

$$\sum_{i \in Q^\alpha} \Pi_i(\boldsymbol{\epsilon}) + p\overline{X}_{Q^\alpha} \geq \sum_{i \in Q^\alpha} \Pi_i^\alpha(\boldsymbol{\epsilon}) + p^\alpha \overline{X}_{Q^\alpha},$$

with a strict inequality if p is different from p^α. The left and right hand sides of the inequality are respectively the sum of the profits of the firms in Q^α plus the revenues from the sale of the permits for the regulation agency in α under merged ETS and under separate ETS respectively: when $p < p^\alpha$ the increase in firms' profits more than compensates the decrease in the ETS revenues (because of firms' adjustment) and when $p > p^\alpha$, the decrease in firms' profits is more than compensated by the increase in the ETS revenues.

Then, Appendix B shows that the gains from merging ETS at the level of an area is proportional to the expectation of the square of the difference in prices effective in this area between the situation with separate ETS and that with merged ETS, which can be decomposed into two non-negative terms. The first term is the square of the difference between expected ETS prices in both situations, $(\mathbb{E}[\widetilde{p}^\alpha] - \mathbb{E}[\widetilde{p}])^2$. Unsurprisingly, the larger the difference in the expected prices between the two situations, the larger the increase in overall profits due to the equalization of marginal costs on average. The second term is the variance of the difference between $b_{Q^\alpha}\widetilde{\epsilon}_{Q^\alpha}$ and $b_Q\widetilde{\epsilon}_Q$, that is: $\mathbb{V}[b_{Q^\alpha}\widetilde{\epsilon}_{Q^\alpha} - b_Q\widetilde{\epsilon}_Q]$. Merging ETS therefore induces no additional welfare gain in the presence of shocks if all firms in Q are subject to a common shock on their marginal abatement cost. But as soon as this is not the case, merging ETS induces a strictly positive additional welfare gain due to a strict improvement in the absorption of shocks, and this gain increases when the global ETS incorporates firms that are subject to shocks on their marginal abatement costs that are less correlated with the shocks affecting firms in the ETS of area α.

Our reasoning has up to now assumed that the tax levels are fixed and equal across

[24]One cannot exclude $\sum_{i \in Q^\alpha} x_i^\alpha < \overline{X}_{Q^\alpha}$. In that case, the price equilibrium p^α is null so the revenues before the merger are null; as for the revenues after the ETS are merged, they are also equal to $p\overline{X}_{Q^\alpha}$: either $p > 0$ and all permits are sold or $p = 0$. Hence the same inequality holds.

the two scenarios. According to the following proposition, this is a valid assumption when the initial tax levels form an equilibrium.

Proposition 6. *Given the scopes (T^α, Q^α) for $\alpha \in \mathcal{W}$, let the policies $(t^\alpha, \overline{X}_{Q^\alpha})$ form a Nash equilibrium as in the previous sub-section. If ETS are merged, the same tax rates form an equilibrium in the game where areas can only choose their tax rates.*

This last proposition shows that, starting from the Nash equilibrium characterized in the previous sub-section, merging ETS does not induce areas to change their specific tax rates thereafter. The argument is as follows. The optimal tax level (best response) in an area depends on the taxes chosen by the other areas and the overall emission level. Since the merger does not affect the emission volumes of the ETS sectors, the best responses are unaffected by the merging, hence the Nash equilibrium is unchanged as well. Note that once ETS are merged, cost efficiency does not hold anymore *within* an area: the local tax rate, which is initially equal to the expected price within the area, will in general differ from the expected price on the merged ETS.

Overall, Propositions 5 and 6 suggest that simply merging ETS is an easy way for separate areas to improve on global efficiency without too much adjustments. Merging ETS benefits all areas, hence it does not make it necessary to implement compensatory transfers across areas. Furthermore, it does not lead areas to change their tax rates when they are initially in equilibrium. Merging ETS is thus a simple and beneficial mechanism, which should be consensual. Some firms, however, may object to the merging of ETS. As we have seen, the firms' profits decrease in an area for the shocks under which the merged ETS price is higher than the separate ETS price. This occurs in particular when the separate ETS price is lower than that of all other areas. This is more likely to occur, the lower the expected price relative to that of others.[25]When the policies are equilibrium policies, the expected prices are equal to the taxes in each area. Thus, the firms in the area in which the tax level is the lowest may object a merger of ETS.

Finally, let us address the question of the design of ETS in anticipation of a merger of the systems. For given scopes and given ETS quotas in each area, Proposition 5 shows that area α can anticipate an additional welfare benefit due to the merger, proportional to the $\mathbb{E}[(\widetilde{p}^\alpha - \widetilde{p})^2]$. If one considers the non-cooperative choice of policies anticipating that ETS will be merged, each area will of course still maintain cost efficiency, but each area will have an incentive to deviate from the equilibrium characterized in Proposition 4 so as to take into account its additional welfare benefit. Therefore, if the expected price on the ETS in area α as predicted by Proposition 4 (i.e. without anticipating the merger) is smaller (resp. larger) than the expected price of the merged ETS, area α has an incentive to design a quota on its ETS larger (resp. smaller) than in Proposition 4 so as to increase its expected benefit from the merger. Strategic anticipation of the

[25]An increase in the expected price does not necessarily imply a decrease in expected profit, as there may be an additional benefit to the merger due to a decrease in the price volatility.

merger of cap-and-trade mechanisms therefore induces an increase in the differences of the regulation framework across areas.

4 Conclusion

This paper has proposed a normative analysis of the global design of environmental regulation in a world plagued with uncertainty, thereby providing a framework to understand when a mixed regulation, relying on both a cap-and-trade mechanism and a tax system, is optimal.

Adopting then a more realistic approach of non-cooperative design of local regulation by local authorities, the paper has characterized the inefficiencies involved and it has provided a strong Pareto-type argument in favor of merging existing ETS, even with different scopes across areas.

Our analysis relies on some key assumptions that have been discussed and that could be relaxed. The assumption of a linear-quadratic setting is obviously strong and it facilitates the characterization of optimal policies and welfare effects. In a general model, the Pareto-improvement obtained by merging ETS would probably cease to be true and only an aggregate (worldwide) welfare improvement would remain. The results still suggest that the issue of compensating transfers may not be critical in a smooth enough world for which the quadratic approximation is acceptable locally.

The assumption of separability of the environmental regulation and the goods markets has also been discussed and can be relaxed. Still, some separability is necessary for our results. Reintroducing a richer interaction between the goods markets and the environmental regulation would invalidate our separate analysis of policy and scope and would surely forbid any Pareto argument for merging ETS. But again, if the interaction is limited, as seems to be the case in the data, the departure from our results should be limited.

There is a last dimension that we neglect in our paper, namely, the political economy considerations behind the design of environmental regulation. We have discussed briefly how firms may oppose a cap-and-trade system and we also have mentioned the distributive effects within each area of the merger of ETS. But a more systematic understanding of the political economy forces at play would certainly enrich our understanding of how climate regulation is shaped and if and how it can evolve.

References

Ambec, S. and Coria, J. (2013): "Prices vs quantities with multiple pollutants," *Journal of Environmental Economics and Management*, 66(1), 123-140.

Anger, N. (2008): "Emissions trading beyond Europe: Linking schemes in a post-Kyoto world," *Energy Economics*, 30(4), 2028-2049.

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

Carlen, B., and A. Hernandez (2013): "Indexing European CArbon Taxes to the EU ETS Permit Price - A Good Idea?", CTS Working Paper 2013-33, Centre for Transport Studies, Stockholm.

Chevallier, J. (2011): "A Model of Carbon Price Interactions with Macroeconomic and Energy Dynamics," *Energy Economics*, 33(6), 1295-1312.

Department of Climate Change, (2008): "Carbon Pollution Reduction Scheme: Australia's Low Pollution Future," White Paper, Commonwealth of Australia, Canberra.

Flachsland, C., Marschinski, R. and O. Edenhofer (2009): "To Link or not to Link: Benefits and Disadvantages of Linking Cap-and-Trade Systems", *Climate Policy*, 9(4), 358-372.

Guesnerie, R. (2010), *Pour une politique climatique globale-Blocage et ouvertures,* Opuscules du CEPREMAP, Editions Rue d'Ulm, Paris.

Heindl, P., P. Wood and F. Jotzo (2014): "Combining International Cap-and-Trade with National Carbon Taxes", CCEP working paper 1418, Australian National University.

Helm, C. (2003): "International Emissions Trading with Endogenous Allowance Choices," *Journal of Public Economics*, 87(12), 2737-2747.

Jotzo, F. and Betz, R. (2009): "Linking the Australian Emissions Trading Scheme," Research Reports 0914, Environmental Economics Research Hub, Crawford School of Public Policy, The Australian National University.

Mace, M. J., Miller, I., Schwarte, C., Anderson, J., Broekhoff, D., Bradley, R., Bowyer, C. and Heilmayr, R. (2008): "Analysis of the Legal and Organizational Issues Arising in Linking the EU Emissions Trading Scheme to Other Existing and Emerging Emissions Trading Schemes," Foundation for International Environmental Law and Development, London.

Martin, R., de Preux, L.B., and Wagner, U.J., (2014): "The Impact of a Carbon Tax on Manufacturing: Evidence from Microdata", *Journal of Public Economics*, 117, 1-14.

Metcalf, G.E. and Weisbachy, D. (2012): "Linking Policies When Tastes Differ: Global Climate Policy in a Heterogeneous World," *Review of Environmental Economics and Policy*, 6(1), 110-129.

Mandell, S. (2008): "Optimal Mix of Emissions Taxes and Cap-and-Trade," *Journal of Environmental Economics and Management*, 56(2), 131-140.

Newell, R. G. and Pizer, W. A. (2008): "Indexed regulation," *Journal of Environmental Economics and management*, 56(3), 221-233.

Nordhaus, W., (2007): "A review of the 'Stern Review' on the Economics of Climate Change", *Journal of Economic Literature*, 45(3), 686-702.

Pizer, W. A. (2002): "Combining price and quantity controls to mitigate global climate change," *Journal of public economics*, 85(3), 409-434.

Ranson, M. and Stavins, R.N. (2012): "Post-Durban Climate Policy Architecture Based on Linkage of Cap-and-Trade Systems", Climate Change and Sustainable Development Series, 43.

Roberts, M. J. and Spence, M. (1976): "Effluent charges and licenses under uncertainty," *Journal of Public Economics*, 5(3), 193-208.

Sterk, W. and J. Kruger (2009): "Establishing a Transatlantic Carbn Market", *Climate Policy*, 9(4), 389-401.

Stern, N., (2007), *The Economics of Climate Change: The Stern Review*, Cambridge University Press.

Talberg, A. and K. Swoboda (2013): "Emissions Trading Schemes around the World", background note, Parliament of Australia.

Tuerk, A., Mehling, M., Flachsland, C. and W. Sterk (2009): "Linking Carbon Markets: Concepts, Case Studies and Pathways", *CLimate Policy*, 9(4), 341-357.

Weitzman, M.L. (1974): "Prices vs. Quantities," *The Review of Economic Studies*, 41(4), 477-491.

Wood, P., Heindl, P., Jotzo, F. and A. Loschel (2013): "Linking Price and Quantity Pollution Controls under Uncertainty", ZEW discussion paper 13-025.

Appendix A: Global normative analysis

First best optimum. To avoid repetition we consider here the general case where there are shocks $\widetilde{\boldsymbol{\epsilon}}$ within an area and outside emissions given by $y + \widetilde{\eta}$. In the case of a worldwide planner, all the shocks are included in $\boldsymbol{\epsilon}$.

Now the first best level of emissions is defined when the planner has full information on the shocks realized in its area and the outside shock $\widetilde{\eta}$. In that case the regulator maximizes social welfare by choosing $x_i(\boldsymbol{\epsilon}, \eta)$, which in our quadratic setting is fully characterized by the FOC:

$$\text{for any } i \in N,\ b_i(z_i + \epsilon_i - x_i) = \nu + A(X_N + y + \eta) \equiv m(\boldsymbol{\epsilon}, \eta).$$

Dividing by b_i for each i, summing up over all $i \in N$, and gathering the terms in X_N, one gets:

$$\begin{aligned} X_N(\boldsymbol{\epsilon}, \eta) &= \frac{b_N z_N - (\nu + A(y + \eta))}{A + b_N} + \frac{b_N}{A + b_N}\epsilon_N & (24) \\ m(\boldsymbol{\epsilon}, \eta) = b_N(z_N + \epsilon_N - X_N(\boldsymbol{\epsilon}, \eta)) &= \frac{b_N(\nu + A(y + \eta + z_N)}{A + b_N} + \frac{A b_N}{A + b_N}\epsilon_N & (25) \\ x_i(\boldsymbol{\epsilon}, \eta) &= z_i + \epsilon_i - \frac{m(\boldsymbol{\epsilon}, \eta)}{b_i} \text{ for all } i \in N & (26) \end{aligned}$$

Let $y = 0$. Absent any uncertainty, i.e. setting $\epsilon_i \equiv 0$ for all i and $\eta \equiv 0$, one obtains the characterization (7) with the value m defined by $m = \frac{b_N(\nu + A z_N)}{A + b_N}$ and then, for any subset S of N, $X_S^* = z_S - \frac{m}{b_S}$. Under uncertainty, writing all variables as deviations from the same variables absent uncertainty, i.e. from the variables with a $*$, we obtain (8).

When $y > 0$, the expressions are simply adjusted by constant terms, replacing ν by $\nu + Ay$ in the quantities with a $*$ without uncertainty.

Proof of Proposition 1. As above, we include the proof in the case of an outside emission $y + \eta$. For any $i \in T$, firm i maximizes its profit net of the tax ($\tau = t$), i.e. it chooses x_i such that $b_i(z_i + \epsilon_i - x_i) = t$ when the shock is ϵ_i; then, $X_T = z_T - \frac{t}{b_T} + \epsilon_T$. For any $i \in Q$, firm i maximizes its profit net of the cost of purchasing the x_i permits ($\tau = p$), and the market for permits clears at a perfectly competitive price p given the realization of uncertainty $\boldsymbol{\epsilon}$:

$$\begin{aligned} x_i &= z_i + \epsilon_i - \frac{p}{b_i}, \\ \sum_{i \in Q} x_i &= \overline{X}_Q. \end{aligned}$$

Summing up all x_i for $i \in Q$, one gets: $p = b_Q(z_Q - \overline{X}_Q + \epsilon_Q)$. Moreover, aggregating over Q, the following holds: $z_i + \epsilon_i - x_i = \frac{b_Q}{b_i}(z_Q + \epsilon_Q - \overline{X}_Q)$.

Under scope (T, Q) and policy $(t, \overline{X}_Q)$, the value of social welfare ex post, when the shocks are $\boldsymbol{\epsilon}$ and η, is then given by:

$$\begin{aligned} W^{T,Q}(\boldsymbol{\epsilon}, \eta, t, \overline{X}_Q) &= \lambda + \sum_{i \in N} \xi_i - \frac{t^2}{2b_T} - \frac{b_Q}{2}(z_Q + \epsilon_Q - \overline{X}_Q)^2 \\ & \quad -\nu(y + \eta + z_T + \epsilon_T - \frac{t}{b_T} - \overline{X}_Q) - \frac{A}{2}(y + \eta + z_T + \epsilon_T - \frac{t}{b_T} - \overline{X}_Q)^2. \end{aligned}$$

The necessary and sufficient FOC for the maximization of the area's social welfare are given by:

$$\begin{aligned} 0 &= \mathbb{E}\left[\frac{\partial W^{T,Q}}{\partial t}(\boldsymbol{\epsilon}, t, \overline{X}_Q)\right] = -\frac{t}{b_T} + \frac{\nu}{b_T} + \frac{A}{b_T}(y + z_T + \overline{X}_Q - \frac{t}{b_T}) \\ 0 &= \mathbb{E}\left[\frac{\partial W^{T,Q}}{\partial \overline{X}_Q}(\boldsymbol{\epsilon}, t, \overline{X}_Q)\right] = b_Q(z_Q - \overline{X}_Q) - \nu - A(y + z_T + \overline{X}_Q - \frac{t}{b_T}). \end{aligned}$$

From these, it follows that:

$$\nu + A(y + z_T + \overline{X}_Q - \frac{t}{b_T}) = t = b_Q(z_Q - \overline{X}_Q).$$

So, $t = m$ and $\overline{X}_Q = X_Q^*$. It follows that $X_N = X_Q^* + z_T - \frac{m}{b_T} + \epsilon_T = X_N^* + \epsilon_T = z_N - \frac{m}{b_N} + \epsilon_T$ and $p = m + b_Q \epsilon_Q$. ∎

Proof of Proposition 2. We first prove the formula (14) which gives a decomposition of the loss in welfare due to the scope relative to the ex-post optimal allocation. Fix $\boldsymbol{\epsilon}$ the vector of all shocks.

Let us define $W^{opt}(\boldsymbol{\epsilon}, \Delta)$ as the maximal welfare when the emissions volume is fixed at $X_N = X_N^* + \Delta$. We surely have $W^{fb}(\boldsymbol{\epsilon}) = W^{opt}(\boldsymbol{\epsilon}, \Delta^{opt}(\boldsymbol{\epsilon}))$ where $\Delta^{opt}(\boldsymbol{\epsilon})$ denotes the optimal emission volume.

For a scope (T, Q) and the associated optimal mixed policy, we know that $X_N = X_N^* + \epsilon_T$ and we write the following decomposition:

$$\begin{aligned} W^{fb}(\boldsymbol{\epsilon}) - W^{T,Q}(\boldsymbol{\epsilon}) &= [W^{fb}(\boldsymbol{\epsilon}) - W^{opt}(\epsilon_T)] + [W^{opt}(\epsilon_T) - W^{T,Q}(\boldsymbol{\epsilon})] \\ &= [W^{opt}(\boldsymbol{\epsilon}, \Delta^{opt}(\boldsymbol{\epsilon})) - W^{opt}(\boldsymbol{\epsilon}, \epsilon_T)] + [W^{opt}(\boldsymbol{\epsilon}, \epsilon_T) - W^{T,Q}(\boldsymbol{\epsilon})] \end{aligned} \tag{27}$$

In (27), the first term inside the square brackets is the loss due to a sub-optimal quantity and the second one is the loss due to the inefficient allocation of this quantity.

Computation of $W^{opt}(\boldsymbol{\epsilon}, \Delta)$. When the emissions volume is fixed, only the allocation of this volume among the firms matters. The optimal allocation is obtained by equalizing of the firms' marginal abatement costs:

$$b_i(z_i + \epsilon_i - x_i^{opt}) = b_N(z_N + \epsilon_N - X_N^* - \Delta) = m + b_N(\epsilon_N - \Delta) \tag{28}$$

The overall ex post welfare associated to the optimal allocation of the emission target $X_N^* + \Delta$ is:

$$W^{opt}(\boldsymbol{\epsilon}, \Delta) = \lambda + \sum_{i \in N} \xi_i - \nu(X_N^* + \Delta) - \frac{A}{2}(X_N^* + \Delta)^2 - \frac{1}{2b_N}[m + b_N(\epsilon_N - \Delta)]^2.$$

Using the fact that $X_N^* = z_N - \frac{m}{b_N}$ and $\nu + AX_N^* = m$, we thus obtain

$$W^{opt}(\boldsymbol{\epsilon}, \Delta) = W^* - m\epsilon_N - \frac{1}{2}\left[(A + b_N)\Delta^2 + b_N\epsilon_N^2 - 2b_N\Delta\epsilon_N\right].$$

Loss due to a sub-optimal emission level. It follows that the optimal amount given $\boldsymbol{\epsilon}$ is $\Delta^{opt}(\boldsymbol{\epsilon}) = \frac{b_N\epsilon_N}{A+b_N}$ and the loss due to an emission Δ that is optimally allocated among the firms is:

$$W^{fb}(\boldsymbol{\epsilon}) - W^{opt}(\boldsymbol{\epsilon}, \Delta) = W^{opt}(\boldsymbol{\epsilon}, \Delta^{opt}(\boldsymbol{\epsilon})) - W^{opt}(\boldsymbol{\epsilon}, \Delta) = \frac{A + b_N}{2}\left[\Delta^{opt}(\boldsymbol{\epsilon}) - \Delta\right]^2.$$

Applying this to $\Delta = \epsilon_T$ yields that the first term of (27) is $\frac{A+b_N}{2}\left(\frac{b_N}{A+b_N}\epsilon_N - \epsilon_T\right)^2$, which is equal to the first term of (14).

Let us now compute the second term of (27), and show that it is given by the second term of (14), which will prove (14):

$$W^{opt}(\boldsymbol{\epsilon}, \epsilon_T) - W^{T,Q}(\boldsymbol{\epsilon}) = \frac{b_Q b_N}{2b_T}\epsilon_Q^2.$$

The difference in welfare is equal to the difference in the abatement costs between the two allocations since the emission volume is identical equal to ϵ_T in both. We use the following lemma, which follows from standard computation.

Lemma 1. *Given $\boldsymbol{\epsilon}$ and the volume $X_N^* + \Delta$, consider the optimal allocation x_i^{opt} and another allocation (x_i) that sums to Δ. The difference in the abatement costs between these allocations satisfies*

$$\sum_{i \in N} \frac{b_i}{2}[(z_i + \epsilon_i - x_i)^2 - \left(z_i + \epsilon_i - x_i^{opt}\right)^2] = \sum_{i \in N} \frac{b_i}{2}\left(x_i^{opt} - x_i\right)^2. \tag{29}$$

Proof of Lemma 1. We have shown that the optimal allocation x_i^{opt} is characterized by the equalization of the firms' marginal abatement costs (28). Using that $z_i + \epsilon_i = x_i^{opt} + c/b_i$ where c is the right hand side of (28), write

$$\begin{aligned}(z_i + \epsilon_i - x_i)^2 - \left(z_i + \epsilon_i - x_i^{opt}\right)^2 &= \left(x_i^{opt} - x_i\right)\left(2z_i + 2\epsilon_i - x_i - x_i^{opt}\right) \\ &= \left(x_i^{opt} - x_i\right)(2c/b_i + x_i^{opt} - x_i).\end{aligned}$$

Multiplying by $b_i/2$ this equation and summing over i, the terms associated to c cancel out since by definition $\sum_{i\in N} x_i = \sum_{i\in N} x_i^{opt}$. We obtain (29). ■

Let us apply (29) to $\Delta = \epsilon_T$ and the allocation associated to the scope (T, Q). This allocation satisfies:

$b_i(z_i + \epsilon_i - x_i) = m$ for i in T and $b_i(z_i + \epsilon_i - x_i) = m + b_Q\epsilon_Q$ for i in Q.

Comparing with (28) and using $\epsilon_N - \epsilon_T = \epsilon_Q$ we obtain:

$b_i\left(x_i^{opt} - x_i\right) = -b_N\epsilon_Q$ for i in T and $b_i\left(x_i^{opt} - x_i\right) = (-b_N + b_Q)\epsilon_Q = \frac{b_N b_Q}{b_T}$ for i in Q.

Plugging these expressions into (29) yields

$$W^{opt}(\boldsymbol{\epsilon}, \epsilon_T) - W^{T,Q}(\boldsymbol{\epsilon}) = b_N^2\left(\frac{1}{b_T} + \frac{b_Q}{b_T^2}\right)\epsilon_Q^2.$$

Since $\frac{1}{b_T} + \frac{b_Q}{b_T^2} = \frac{b_Q}{b_T}\left[\frac{1}{b_Q} + \frac{1}{b_T}\right] = \frac{b_Q}{b_T}\frac{1}{b_N}$ we finally obtain $W^{opt}(\epsilon_T) - W^{T,Q}(\boldsymbol{\epsilon}) = \frac{1}{2} b_N \frac{b_Q}{b_T}\epsilon_Q^2$.

End of the Proof of Proposition 3. It has been proved in the text which uniform system is the best one and it follows that $W^{fb} - W^{unif} = \frac{A+b_N}{2}\mathbb{V}[\frac{\mathcal{A}}{A+b_N}\epsilon_N]$. Let us now prove expression (16), calculating the difference between the expression of $W^{fb} - W^{T,Q}$, given by (15), and the previously obtained expression for $W^{fb} - W^{unif}$:

$$W^{unif} - W^{T,Q} = \frac{1}{2}(A + b_N)\left\{\mathbb{V}\left[\frac{b_N}{A + b_N}\epsilon_N - \epsilon_T\right] - \mathbb{V}\left[\frac{\mathcal{A}}{A + b_N}\epsilon_N\right]\right\} + \frac{b_Q b_N}{2b_T}\mathbb{V}[\epsilon_Q] \quad (30)$$

To compute the difference in variances, we decompose $\epsilon_N = \epsilon_T + \epsilon_Q$ and we develop the terms:

$$\begin{aligned}\mathbb{V}\left[\frac{b_N}{A + b_N}\epsilon_N - \epsilon_T\right] &- \mathbb{V}\left[\frac{\mathcal{A}}{A + b_N}\epsilon_N\right] = \frac{1}{(A + b_N)^2}\left[\mathbb{V}[b_N\epsilon_Q - A\epsilon_T] - \mathbb{V}[\mathcal{A}\epsilon_Q + \mathcal{A}\epsilon_T]\right] \\ &= \frac{1}{(A + b_N)^2}\left\{(A^2 - \mathcal{A}^2)\mathbb{V}[\epsilon_T] - 2(Ab_N + \mathcal{A}^2)\ \mathrm{cov}(\epsilon_T, \epsilon_Q) + (b_N^2 - \mathcal{A}^2)\mathbb{V}[\epsilon_Q]\right\}\end{aligned}$$

Observe that $A^2 - \mathcal{A}^2$ is null for $A = \mathcal{A}$ and equal to $A^2 - b_N^2 = (A - b_N)(A + b_N)$ for $\mathcal{A} = b_N$, hence $A^2 - \mathcal{A}^2 = (A - \mathcal{A})(A + b_N)$. Similarly $Ab_N + \mathcal{A}^2$ is either equal to $A(b_N + A)$ for $\mathcal{A} = b_N$ and to $(A + b_N)b_N$ for $\mathcal{A} = A$, which implies $Ab_N + \mathcal{A}^2 = (A + b_N)\mathcal{A}$ and $b_N^2 - \mathcal{A}^2 = (b_N - \mathcal{A})(A + b_N)$. Using these in the difference of variances and plugging it into (30) yields (16).

Proof of Proposition 2. A firm strictly prefers the ETS to the tax if and only if $\mathrm{cov}(b_Q\widetilde{\epsilon}_Q, b_Q\widetilde{\epsilon}_Q - 2b_i\widetilde{\epsilon}_i) > 0$. Assume by contradiction that it is satisfied for each i in Q. Dividing by b_i and summing over i in Q yields:

$$\mathrm{cov}(b_Q\widetilde{\epsilon}_Q, \sum_{i\in Q}\frac{b_Q}{b_i}\widetilde{\epsilon}_Q - 2\sum_{i\in Q}\widetilde{\epsilon}_i) > 0 \Leftrightarrow\ \mathrm{cov}(b_Q\widetilde{\epsilon}_Q, -\widetilde{\epsilon}_Q) > 0,$$

which is impossible.

Proof of (a). Under a common shock $b_i\epsilon_i = b_Q\epsilon_Q = \theta$ for all i and the result follows.

Proof of (b). Let shocks have identical variance σ^2, a correlation coefficient ρ and

$b_i = b$ for all i. Then $b_Q = \frac{b}{q}$ where q denotes the number of firms included in the ETS. Let i in Q. Up to the factor σ^2,

$$\mathrm{var}(\widetilde{\epsilon}_Q) = q[1 + (q-1)\rho] \text{ and } \mathrm{cov}(b_Q\widetilde{\epsilon}_Q, b_i\widetilde{\epsilon}_i) = \frac{b^2}{q}[1 + (q-1)\rho].$$

This gives:

$$\mathrm{cov}(b_Q\widetilde{\epsilon}_Q, b_Q\widetilde{\epsilon}_Q - 2b_i\widetilde{\epsilon}_i) = -\frac{b^2}{q}[1 + (q-1)\rho].$$

As $\mathrm{var}(\widetilde{\epsilon}_Q) = q[1 + (q-1)\rho]$ can only be non-negative, this proves the result. ■

Appendix B: A world with several areas

Worldwide optimum. Using the decomposition of A and ν, it follows immediately:

$$\left\{1 + (\sum_{\alpha\in\mathcal{W}} A^\alpha)\left(\sum_{\alpha\in\mathcal{W}} \frac{1}{b_{N^\alpha}}\right)\right\} X^* = z_N - \left(\sum_{\alpha\in\mathcal{W}} \nu^\alpha\right)\left(\sum_{\alpha\in\mathcal{W}} \frac{1}{b_{N^\alpha}}\right).$$

Proof of Proposition 4. For given scopes in each area, the general analysis applies at the level of each area α so as to find out this area's best response policy, taking the other areas' random level of emissions $Y = X^{-\alpha}$ as given.

Let $m^\alpha(y) \equiv \frac{b_{N^\alpha}(\nu^\alpha + A^\alpha y + A^\alpha z_{N^\alpha})}{A^\alpha + b_{N^\alpha}}$. Then, the equilibrium tax rate and the optimal aggregate quota in area α given the aggregate emissions volume in the other area $X^{-\alpha}$ are determined by:

$$\begin{aligned} t^\alpha &= m^\alpha(\mathbb{E}[X^{-\alpha}]), \\ \overline{X}_{Q^\alpha} &= z_{Q^\alpha} - \frac{m^\alpha(\mathbb{E}[X^{-\alpha}])}{b_{Q^\alpha}}. \end{aligned}$$

From these, it follows that $p^\alpha = m^\alpha(\mathbb{E}[X^{-\alpha}]) + b_{Q^\alpha}\epsilon_{Q^\alpha}$ is the equilibrium price on area α's ETS market and $X_{T^\alpha} = z_{T^\alpha} - \frac{m^\alpha(\mathbb{E}[X^{-\alpha}])}{b_{T^\alpha}} + \epsilon_{T^\alpha}$. Finally, in terms of the area's aggregate emissions, as a best response:

$$\mathbb{E}[X^\alpha] = z_{N^\alpha} - \frac{m^\alpha(\mathbb{E}[X^{-\alpha}])}{b_{N^\alpha}}.$$

Since the best response in terms of $\mathbb{E}[X^\alpha]$ are decreasing with slope of absolute value less than one, a sufficient condition for an interior equilibrium is (by analogy with a Cournot model): for $\alpha \in \{\alpha, \beta\}$

$$z_{N^\alpha} - \frac{m^\alpha(0)}{b_{N^\alpha}} \leq \frac{b_{N^{-\alpha}} z_{N^{-\alpha}} - \nu^{-\alpha}}{A^{-\alpha}}.$$

It is then immediate to check that $m^\alpha(\mathbb{E}[X^{-\alpha}]) = \nu^\alpha + A^\alpha\mathbb{E}[X^\alpha + X^{-\alpha}] = b_i(z_i - \mathbb{E}[x_i])$

for any $i \in N^\alpha$, and the usual manipulation yields:

$$\left\{1 + \left(\sum_{\alpha\in\mathcal{W}} \frac{A^\alpha}{b_{N^\alpha}}\right)\right\} \mathbb{E}[X^\alpha + X^{-\alpha}] = z_N - \left(\sum_{\alpha\in\mathcal{W}} \frac{\nu^\alpha}{b_{N^\alpha}}\right).$$

The comparison with the similar expression for X^* yields unsurprisingly that the (Nash) equilibrium in policy mix yields expected emissions volumes larger than X^* as the equilibrium situation does not enable to internalize the externalities worldwide, hence too much emission on the aggregate. Moreover,

$$t^\alpha = \nu^\alpha + A^\alpha \mathbb{E}[X^\alpha + X^{-\alpha}] = \nu^\alpha + A^\alpha X^* + A^\alpha(\mathbb{E}[X] - X^*) \equiv m^\alpha + A^\alpha(\mathbb{E}[X] - X^*) > m^\alpha$$

$$\overline{X}_{Q^\alpha} = \frac{\mu_{Q^\alpha} - t^\alpha}{b_{Q^\alpha}} = \frac{\mu_{Q^\alpha} - m}{b_{Q^\alpha}} + \frac{m - t^\alpha}{b_{Q^\alpha}} = X^*_{Q^\alpha} + \frac{m - t^\alpha}{b_{Q^\alpha}}$$

so that $\overline{X}_{Q^\alpha} > X^*_{Q^\alpha} \Leftrightarrow m > t^\alpha$.

Finally, let us investigate whether in equilibrium it is possible that $t^\alpha > m$. Suppose here that there are only two areas, α and β, and that both A^α and ν^α are scaled up by a factor k. Then:

$$\begin{aligned} t^\alpha > m \quad &\Leftrightarrow \quad k(\nu^\alpha + A^\alpha \mathbb{E}[X]) > k(\nu^\alpha + A^\alpha X^*) + (\nu^\beta + A^\beta X^*) \\ &\Leftrightarrow \quad A^\alpha(\mathbb{E}[X] - X^*) > \frac{1}{k}(\nu^\beta + A^\beta X^*). \end{aligned}$$

When $1/k$ becomes negligible, the difference $\mathbb{E}[X] - X^*$ is bounded from below by a positive value while the RHS of the above inequality becomes very small. Therefore, when $1/k$ becomes negligible, area α compensates for area β's lax policy. ■

Proof of formula (23). We extend the computation made for proving formula (14). Consider an area, and let η be the realized shock within the area and $y + \eta$ be the outside emission level. Now the first best level of emissions for the area when it has full information on the shocks realized in its area and the outside shock η is given by (24). Under a scope (T, Q), the emission level by the firms within the area is still given by $X^*_T + \epsilon_T$, where X^*_T accounts for the expected outside emission level y, as follows from the proof of Proposition 1 and (26). Thus firms' emission levels only depend on the shocks $\boldsymbol{\epsilon}$ and total emissions are $X^*_T + y\eta + \epsilon_T$.

The loss in welfare can still be decomposed into two terms. The second term is unchanged, as it comes from the inefficiencies within the firms in the country. The first term corresponds to the loss due to an inefficient emission level and is equal here to $\mathbb{E}[W^{opt}(\widetilde{\boldsymbol{\epsilon}}, \widetilde{\eta}, \Delta^{opt}(\widetilde{\boldsymbol{\epsilon}}, \widetilde{\eta})) - W^{opt}(\widetilde{\boldsymbol{\epsilon}}, \widetilde{\eta}, \widetilde{\epsilon}_T)]$, in which $\Delta^{opt}(\boldsymbol{\epsilon}, \eta)$ denotes the optimal quantity to be emitted within the country, given $\boldsymbol{\epsilon}$ and η.

We proceed in the same way as in the case of a single area. Denoting by Δ the emissions level in the country, we have that the overall ex post welfare associated to the optimal allocation of the emission target $X_N^* + \Delta$ is:

$$W^{opt}(\boldsymbol{\epsilon}, \eta, \Delta) = \lambda + \sum_{i \in N} \xi_i - \nu(X_N^* + \Delta + \eta) - \frac{A}{2}(X_N^* + \Delta + \eta)^2 - \frac{1}{2b_N}[m + b_N(\epsilon_N - \Delta)]^2.$$

Using the fact that $X_N^* = z_N - \frac{m}{b_N}$ and $\nu + AX_N^* = m$, we thus obtain

$$W^{opt}(\boldsymbol{\epsilon}, \eta, \Delta) = W^*(\epsilon_N, \eta) - \frac{A}{2}\eta_A^2 - \frac{1}{2}\left[(A + b_N)\Delta^2 + b_N\epsilon_N^2 - 2b_N + 2A\Delta\epsilon_N + 2A\Delta\eta\right].$$

where $W^*(\epsilon_N, \eta)$ does not depend on Δ ($=W^* - m\epsilon_N - (\nu + AX_N^*)\eta$). It follows that the optimal amount given $\boldsymbol{\epsilon}, \eta$ is $\Delta^{opt}(\boldsymbol{\epsilon}, \eta) = \frac{b_N\epsilon_N - A\eta}{A + b_N}$. As $W^{fb}(\boldsymbol{\epsilon}, \eta,) = W^{opt}(\boldsymbol{\epsilon}, \eta, \Delta^{opt}(\boldsymbol{\epsilon}, \eta))$, the loss due to an (optimally allocated among the firms) emission Δ is:

$$W^{fb}(\boldsymbol{\epsilon}, \eta,) - W^{opt}(\boldsymbol{\epsilon}, \Delta) = \frac{A + b_N}{2}\left[\Delta^{opt}(\boldsymbol{\epsilon}, \eta) - \Delta\right]^2.$$

Applying this to $\Delta = \epsilon_T$ and taking expectation over the shocks yields that the first term of (23) is equal to $\frac{A+b_N}{2}\left(\frac{b_N}{A+b_N}\epsilon_N - \epsilon_T\right)^2$. ■

Proof of Proposition 5. We compare the situation with merged ETS with the situation with separate ETS, using the notations introduced in the text.

First, note that plugging the value of x_i^α and x_i into the expressions for the profits, it comes: $\Pi_i^\alpha(\boldsymbol{\epsilon}) = \xi_i - \frac{(p^\alpha)^2}{2b_i} - p^\alpha x_i^\alpha$ and $\Pi_i(\boldsymbol{\epsilon}) = \xi_i - \frac{(p)^2}{2b_i} - px_i$. Summing up over $i \in Q^\alpha$, adding up the revenues from the sale of permits and noticing that $\sum_{i \in Q^\alpha} x_i^\alpha = \overline{X}_{Q^\alpha}$, it comes:

$$\begin{aligned}
\sum_{i \in Q^\alpha} \Pi_i^\alpha(\boldsymbol{\epsilon}) + p^\alpha \overline{X}_{Q^\alpha} &= \sum_{i \in Q^\alpha} \xi_i - \frac{(p^\alpha)^2}{2b_{Q^\alpha}} \\
\sum_{i \in Q^\alpha} \Pi_i(\boldsymbol{\epsilon}) + p\overline{X}_{Q^\alpha} &= \sum_{i \in Q^\alpha} \xi_i - \frac{p^2}{2b_{Q^\alpha}} + p\sum_{i \in Q^\alpha}(x_i^\alpha - x_i).
\end{aligned}$$

Using the fact that $x_i^\alpha - x_i = \frac{p - p^\alpha}{b_i}$, the difference in social welfare for area α between the situation with merged ETS and that with separate ETS can be written:

$$\frac{(p^\alpha)^2 - p^2}{2b_{Q^\alpha}} + p\frac{(p - p^\alpha)}{b_{Q^\alpha}} = \frac{(p^\alpha - p)^2}{2b_{Q^\alpha}}.$$

Proposition 5 follows. ■

Proof of Proposition 6. Let us consider that the climate policies are the equilibrium ones and consider a simple game between the areas where each of them chooses non-cooperatively its tax rate t^α. In this game with separate ETS, using Appendix 1, it is immediate that the best response tax rate in area α is still given by (21) :

$$t^\alpha = \nu^\alpha + A^\alpha\mathbb{E}[X^\alpha + X^{-\alpha}].$$

As this expression only depends upon the tax rates and the sum of the quotas, the same expression also determines the best response tax rate in area α under merged ETS with the sum of quotas. Hence, the conclusion. $\blacksquare$