



On the Effects of Linking Cap-and-Trade Systems for CO₂ Emissions

Bjart Holtsmark¹ · Martin L. Weitzman²

Accepted: 7 January 2020 / Published online: 23 January 2020
© The Author(s) 2020

Abstract

Linkage of national cap-and-trade systems is typically advocated by economists on a general analogy with the beneficial linkage of free-trade areas and on the specific grounds that linkage will ensure cost effectiveness among the linked jurisdictions. The paper analyses the less obvious effects of linkage with the bottom–up approach of the Paris Agreement where each country sets its nationally determined contribution for its own carbon dioxide (CO₂) emissions. An appropriate and widely accepted specification for the damages of CO₂ emissions within a relatively short (say 5–10 year) period is that marginal damages for each jurisdiction are constant (although they can differ among jurisdictions). With this defensible assumption, the analysis is significantly clarified and yields simple closed-form expressions for all CO₂ permit prices. Some implications for linked and unlinked voluntary CO₂ cap-and-trade systems are derived and discussed. A numerical example illustrates the results.

Keywords Cap and trade · Climate change · Linkage · Paris agreement · Pollution

Very sadly, Martin L. Weitzman died on 27 August 2019, before this paper's section 6 was completely finalised. The co-author is deeply grateful for having had the opportunity to work with this great economist and researcher. For helpful constructive comments, but without implicating them in errors, omissions, or interpretations, the authors thank Katinka Holtsmark, Ulrike Kornek, Gilbert Metcalf, Antony Millner, Cedric Philibert, Simon Quemin, Will Rafey, Joseph Shapiro, Luca Taschini, Hidemichi Yonezawa, as well as two anonymous referees.

✉ Bjart Holtsmark
bjart.holtsmark@ssb.no

Martin L. Weitzman
mweitzman@harvard.edu

¹ Research Department, Statistics Norway, Oslo, Norway

² Department of Economics, Harvard University, Cambridge, USA

1 Introduction

Abatement of carbon dioxide (CO₂) emissions is today's premier global public good.¹ It is difficult enough to resolve a local public goods problem within a jurisdiction having effective governance with an ability to levy payments. But when the problem is international in scope, and where there is no overarching top-down international governance structure, it can render a global public goods problem virtually unsolvable.

The key issue here is the notorious free-rider problem. Everyone wants to free ride off the contributions of others. A jurisdiction bears the full costs of its abatement, but it only reaps a fraction of the global benefits. The result is a non-cooperative selfish equilibrium where everyone abates far less than would be socially desirable in a cooperative solution. The key issue is that it takes a strong government to enforce a socially desirable cooperative solution. In the global arena there is no such strong international government with powers to assign CO₂ emissions targets and enforce penalties for non-compliance.

Reduced to its core essence, the Paris Agreement (UNFCCC 2015) is strictly a bottom-up voluntary agreement based on a periodically repeated 'pledge and review' process. Just before a performance period, at the end of the previous performance period, each country volunteers a 'nationally determined contribution' (NDC) for its own CO₂ emissions. After a performance period (5 years), actual emissions should be reported but there are no penalties for a country not complying with its own volunteered NDC. In this sense the Paris Agreement is doubly voluntary: The self-announced pledges are strictly voluntary in the first place, and compliance with the previous self-announced pledges in the performance period is also strictly voluntary. The Agreement talks about developed countries aiding developing countries with financial support for 'sustainability' based on climate mitigation and adaptation, but the cash flows have thus far been meager.

Not surprisingly, there has been broad take-up of such a strictly voluntary agreement. Before the U.S. dropped out, the Paris Agreement nominally covered countries accounting for some 97 per cent of world CO₂ emissions (see e.g. Chen 2017). There is widespread acknowledgement that the highly under-ambitious NDCs actually named are not nearly enough to keep global warming on a track below the stated goal of no more than a world-wide average temperature increase of 2 °C (see e.g. Levin and Fransen 2015).

On the positive side, the Paris Agreement has highlighted the importance to the international community of dealing with climate change. And it encourages credibly transparent standards of reporting, monitoring, and verification by each participating country, which is a necessary first step for any accord. The Paris Agreement also contains a commitment for countries to pledge, review, and re-pledge new intended NDCs periodically (every 5 years), hopefully inspiring ever-greater levels of NDC ambition over time (Paragraphs 4.8–4.9).

What is the underlying 'model' of human behavior that might allow the Paris Agreement to be seen as a step toward a resolution of the climate-change externality? There appears to be an implicit assumption here that CO₂ polluters will significantly drive down their emissions voluntarily based largely on altruism and 'blame and shame' from others, without any top-down setting of emissions targets or enforcement of penalties for non-compliance. If only everyone followed the full golden rule, the global-warming problem could be solved. The Paris Agreement might then be seen as a first tentative step toward

¹ For expositional simplicity we pretend that carbon dioxide (CO₂) is the only greenhouse gas (GHG). CO₂ is by a wide margin the most important GHG, but it is not the only GHG.

demonstrating the spirit of golden-rule-like behavior, which might hopefully inspire further steps toward even more golden-rule-like behavior by inspiring ever-more-ambitious NDC targets in a virtuous circle.

There seems to be little question but that some jurisdictions throughout the world have gone beyond the most narrow definition of pure self-interest in proposing relatively more ambitious emissions targets, even if this level of ambition still falls well short of full golden-rule behavior. Altruism may thus help somewhat, and is to be encouraged, although it remains to see whether and to what extent ambitious emission targets will be met. Generally, most economists would express at least partial skepticism about the golden-rule model because it is typically difficult to resolve free-rider problems by altruism alone. This paper goes to the opposite extreme of altruistic behavior by examining the consequences of a model of pure self interest.

A jurisdictional cap-and-trade system assigns allowances of CO₂ caps to emitters within a jurisdiction and then allows (or even encourages) internal free trade in permits. Total emissions of CO₂ must equal the sum of all allocated caps. Regulators of a cap-and-trade system can thus control the total amount of CO₂ emissions within their jurisdiction by controlling the sum of all allocated caps. As economists have long emphasized, a cap-and-trade system is cost effective because it minimizes total abatement costs for each chosen level of total CO₂ emissions via ensuring that every emitter in the jurisdiction sets its marginal cost of abatement equal to the equilibrium price of permits.

Economists have typically advocated *linkage* of different cap-and-trade jurisdictions by rough analogy with the beneficial linkage of free-trade areas and on the more specific basis that this will ensure cost effectiveness among the linked jurisdictions.² An unlinked cap-and-trade system guarantees cost effectiveness only within its own jurisdiction. A linked cap-and-trade system goes further by also ensuring cost effectiveness among the linked jurisdictions taken as a whole.

The argument for linkage based on cost effectiveness might make sense when there is an overall top-down governance structure that can force Pareto-improving side payments among the linked participants. But absent such a powerful overall governance structure, this cost-effectiveness argument in a strictly voluntary cap-and-trade system loses its force. Cap-and-trade jurisdictions that have linked their cap-and-trade systems will issue their own voluntary caps on the basis of a host of domestic political-economy considerations including, prominently, self interest. Narrow self interest, which is being modeled in this paper as if it were the sole motivation, will cause linked jurisdictions to pay relatively little attention to what is for them a strictly hypothetical argument about minimization of overall compliance costs. Be that as it may, there is a widespread feeling among most economists that linking cap-and-trade jurisdictions is a good idea.

The generally favorable attitude toward linking has found its way into the Paris Agreement. Paragraph 6.2 of the agreement outlines a framework for recognizing traded obligations (called ‘internationally transferred mitigation outcomes’) so that double counting is avoided because a party to the agreement is allowed to include traded reductions undertaken by another party to count toward the first party’s NDC. Paragraph 6.3 states that: ‘The use of internationally transferred mitigation outcomes to achieve nationally determined contributions under this Agreement shall be voluntary and authorized by

² See, for example, the extremely comprehensive article of Mehling et al. (2017), and the numerous further references they cite.

participating parties'. The inclusion of articles 6.2 and 6.3 opens the door to linking cap-and-trade systems (or, indeed, any market-based mechanisms).

While linking may give some jurisdictions incentives to choose more ambitious caps, it could also give other jurisdictions incentives to choose less ambitious caps. It is a seeming paradox that cap-and-trade among the parties to the Paris Agreement might lead to even higher emissions. If linking cap-and-trade systems does little more than replace one non-cooperative equilibrium with another, it will still be a far cry from the more cooperative outcome that might accompany a genuine international governance structure.

The insight that linking voluntary cap-and-trade systems may lead to higher levels of pollution emissions is not new.³ The most complete rigorous analysis of this possibility is the pioneering work of Helm (2003), who models both unlinked and linked cap-and-trade systems as a non-cooperative Nash equilibrium among self-interested countries. With a fairly general treatment of environmental pollution damages, he finds that overall effects on total emissions are ambiguous and he derives moderately complicated conditions for when pollution is increased or decreased by linking. He addresses some additional results with the use of quadratic abatement costs functions.

In the present analysis all theoretical results are based on quite general assumptions about abatement cost functions. In the theoretical analysis we found it desirable to stick to this assumption because there is lack of an empirical basis for the assumption that quadratic cost functions represent a realistic approximation. However, an appropriate and widely accepted specification for the damages of CO₂ emissions within a relatively short (5–10 year) period is that marginal damages are constant for each jurisdiction (although they can differ among jurisdictions). With this defensible assumption, which the existing literature has not yet focused upon, the analysis is significantly clarified and yields simple closed-form expressions for all (linked and unlinked) CO₂ permit prices. Some sharp insights are then available. How a linked jurisdiction sets its voluntary caps relative to actual emissions (and whether the jurisdiction buys or sells CO₂ permits) is fully characterized by a simple linear proportionality condition that depends only on the difference between the jurisdiction's marginal damages and the average marginal damages of the entire linked system. Some implications for linked and unlinked voluntary CO₂ cap-and-trade systems are derived and discussed. The theoretical analysis and its implications are illustrated with a numerical example. In this part of the paper we have specified the abatement cost functions as quadratic while the marginal damages are constant (as in the theoretical parts).

2 The Model

The emphasis in the model of this paper is on clarity of exposition and the appealing simplicity of clean crisp analytical results. Hopefully the model embodies enough of 'reality' to give some useful insights on an important issue.

Let there be a total of n (≥ 2) jurisdictions. Throughout this paper we economize on notation by not redundantly pointing out that index i always runs from $i = 1, 2, \dots, n$ or that index j always runs from $j = 1, 2, \dots, n$. Henceforth it is understood that i (or j) refer to one of the n jurisdictions under consideration.

³ Early preliminary intimations of this tendency are expressed in Bohm (1992), Eyckmans and Proost (1996), and Krishna and Tan (1999). This issue is also discussed later in Green et al. (2014) and in Mehling et al. (2017).

For each i , the marginal damage of emissions within a pledge-and-review cycle is given as d_i . The marginal damages curve is thus assumed here to be flat in emissions flows. This assumption, which is standard in the climate-change literature, is appropriate for CO₂ emissions because it is the stock of accumulated CO₂ that does the damage and the relatively small flow of CO₂ emissions within, say, a 5–10 year period has an effectively linear impact on the overall stock of atmospheric CO₂. Let e_i represent the emissions flow of jurisdiction i . Let $D_i(E)$ denote the total damage to jurisdiction i of total emissions $E \equiv \sum e_i$ and let $D'_i(E)$ represent the marginal damage to jurisdiction i of total emissions E . Then we are assuming here that

$$D'_i(E) = d_i, \tag{1}$$

and the $\{d_i\}$ coefficients thus provide an unambiguous ordering of the marginal damages of emissions among the n cap-and-trade jurisdictions.⁴ The constancy of marginal damages for each jurisdiction, which is a natural assumption for CO₂, allows for a simplification of results, which seemingly has not been taken full advantage of in this literature.

On the abatement cost side, let $C_i(e_i)$ denote the cost to jurisdiction i of emissions e_i , where $C'_i(e_i) < 0$ and $C''_i(e_i) > 0$. Note, importantly, that the marginal cost of *abatement* for jurisdiction i is *minus* $C'_i(e_i)$.

Let p be an exogenously imposed tax-price on emissions (it does not matter what is the source of the tax-price p , so long as it is perceived as exogenous). Let e_i here represent the emissions quantity reaction of jurisdiction i to the tax-price p . The functional relationship between e_i and p is given by the condition that marginal abatement cost equals price, or

$$p = -C'_i(e_i). \tag{2}$$

Let the function $e_i(p)$ represent the *inverse* of the marginal cost of abatement function $-C'_i(e_i)$ in (2). We can then write and conceptualize, whenever appropriate to the context, that

$$e_i = e_i(p), \tag{3}$$

where $e'_i(p) < 0$ because $C''_i(e_i) > 0$.

The two simple specifications (1) (which is very specific to CO₂), and (3) (which is entirely general for any cost functions $\{C_i(e_i)\}$ with $C''_i(e_i) > 0$) constitute the analytical framework for the study of unlinked and linked voluntary cap-and-trade systems investigated in this paper.⁵

⁴ Note that, other things being equal, jurisdictions with higher (lower) populations will tend to have larger (smaller) marginal damages of total emissions.

⁵ We can only hope in this paper that, as is often the case in economic theory, an analytically-tractable flow model, standing in for a more complicated stock-flow situation, is capable of offering some useful insights. We have in mind here a pledge-and-review cycle of maybe 5–10 years or so, which may be short enough to justify the model specification here.

3 Unlinked Cap-and-Trade Systems

Independent unlinked jurisdiction i seeks to minimize over e_i the expression $D_i(E) + C_i(e_i)$. Let the solution be denoted \hat{e}_i . Then \hat{e}_i satisfies the first-order condition $D'_i(\sum \hat{e}_j) = -C'_i(\hat{e}_i)$. Substituting from (1) and (3), this first-order condition translates into

$$\hat{e}_i = e_i(d_i), \quad (4)$$

with the corresponding autarchic-internal cap-and-trade emissions price being

$$\hat{p}_i = d_i. \quad (5)$$

Thus, with constant marginal damages, in the unlinked case of this model it turns out that there is no strategic interaction among the n jurisdictions.

The total emissions of all jurisdictions, denoted \hat{E} , is then

$$\hat{E} \equiv \sum \hat{e}_i = \sum e_i(d_i). \quad (6)$$

The free-riding voluntary autarchic emissions levels $\{\hat{e}_i\}$ do not, by a wide margin, represent socially optimal emission levels. The socially optimal level of e_i , denoted e_i^* , satisfies the Lindahl–Samuelson public-goods condition $\sum d_j = -C'_i(e_i^*)$ for all i , whose inverse is $e_i^* = e_i(\sum d_j)$ and which clearly represents a lower level of emissions than \hat{e}_i given by Eq. (4).

The total socially optimal emissions level of all jurisdictions is then

$$E^* \equiv \sum e_i^* = \sum_i e_i \left(\sum_j d_j \right), \quad (7)$$

which is clearly lower than the free-riding total emissions \hat{E} given by Eq. (6). The corresponding uniform shadow price of socially optimal emissions (that also ensures cost effectiveness) is $p^* = \sum d_i$, which is clearly higher than the average unlinked voluntary cap-and-trade price $\sum \hat{p}_i/n = \sum d_i/n$. However, as an extension of the argument about cost effectiveness given in the Introduction, socially optimal levels of CO₂ emissions or the socially optimal shadow emissions price of CO₂ are largely irrelevant for strictly voluntary cap-and-trade systems with no overarching top–down governance structure that can determine the initial allocation of CO₂ caps and penalize non-compliance. Absent such a powerful overall governance structure, the socially optimal solution in a strictly voluntary cap-and-trade system loses much of its rationale.

4 Linked Cap-and-Trade Systems

Suppose that the n jurisdictions have been persuaded to link their cap-and-trade systems. We now investigate the steady-state Nash-equilibrium outcome of such linking. This case presents far more of an analytical challenge than the case of unlinked cap-and-trade systems due to the strategic interaction among linked jurisdictions.

Note that even though we presume that all jurisdictions play individually and non-cooperatively against all other jurisdictions, some aspects of the rules of the game must be agreed upon beforehand. In particular, property rights, which are traded on the permit

market, must be enforced. A jurisdiction must reveal its post-cap-and-trade permits to ensure consistency with its actual post-cap-and-trade emissions. (This is the meaning of Paragraph 6.2 of the Paris Agreement, previously referred to.)

Let the *actual* post-cap-and-trade emissions of jurisdiction i be denoted e_i^a . For any given exogenously-imposed CO₂ permit (or allowance) price P , the actual-emissions reaction of jurisdiction i is obtained by the condition $-C'_i(e_i(P)) = P$. This condition yields the same inverse-function formula as (3), rewritten here for emphasis as

$$e_i^a = e_i(P), \tag{8}$$

where $e'_i(P) < 0$.

Let $E^a \equiv \sum e_i^a$ represent *total actual* post-cap-and-trade emissions. We now seek to find the equilibrium permit price for the linked cap-and-trade system as a function of total actual emissions, denoted $P(E^a)$. Looking at actual emissions and adding up expression (8) over all i , we obtain $E^a(P) = \sum e_i(P)$, which, when inverted, yields the basic equation for the equilibrium permit price as a function of total actual emissions, namely $P(E^a)$.

Next, let e_i^c represent the emissions permits or caps issued voluntarily by jurisdiction i . The superscript ‘ c ’ stands for *cap* (and also for *control* variable). Define $E^c \equiv \sum e_i^c$ to be the total voluntary emissions caps or permits, and note that in equilibrium $E^c = E^a$. Let \tilde{e}_i^c represent the Nash-equilibrium self-interested number of voluntary emissions permits issued by jurisdiction i , contingent on Nash-equilibrium self-interested voluntary emissions permits \tilde{e}_j^c for all other jurisdictions $j \neq i$. Then \tilde{e}_i^c must *maximize* over all possible voluntary emissions caps e_i^c the expression

$$\left\{ P \left(e_i^c + \sum_{j \neq i} \tilde{e}_j^c \right) \cdot (e_i^c - e_i^a) \right\} - \left(D_i \left(e_i^c + \sum_{j \neq i} \tilde{e}_j^c \right) + C_i(e_i^a) \right), \tag{9}$$

where we understand e_i^a in expression (9) as being some implicit function of e_i^c .

The second term of (9), in round brackets, represents the loss of welfare to jurisdiction i from emissions damages and costs.

Let us examine more closely the important expression within the curly brackets of (9). If $e_i^a > e_i^c$, then jurisdiction i is obliged to *buy* from other jurisdictions $(e_i^a - e_i^c)$ emissions permits, costing it a cash *loss* of $P \cdot (e_i^a - e_i^c)$, which renders the expression in the curly brackets of (9) negative, reflecting cash outflows out of jurisdiction i . If $e_i^a < e_i^c$, then jurisdiction i can *sell* to other jurisdictions $(e_i^c - e_i^a)$ emissions permits, earning it a cash *revenue* of $P \cdot (e_i^c - e_i^a)$, which renders the expression in the curly brackets of (9) positive, reflecting cash inflows into jurisdiction i . To summarize here, the expression in the curly brackets of (9) exactly equals the *net* cash flow into jurisdiction i from inter-jurisdictional tradable permits.

The Nash-equilibrium linked cap-and-trade actual emissions of jurisdiction i is denoted \tilde{e}_i^a . Taking derivatives with respect to e_i^c and making use of (1), the first-order condition for \tilde{e}_i^c to maximize over all e_i^c expression (9) in a self-interested Nash equilibrium is

$$\left[\frac{\partial}{\partial e_i^c} \left\{ P \left(e_i^c + \sum_{j \neq i} \tilde{e}_j^c \right) \cdot (e_i^c - e_i^a) \right\} \right]_{e_i^c = \tilde{e}_i^c; e_i^a = \tilde{e}_i^a} - \left(d_i + C'_i(\tilde{e}_i^a) \frac{\partial e_i^a}{\partial e_i^c} \right) = 0. \tag{10}$$

Let $\tilde{E} (= \sum \tilde{e}_i^c = \sum \tilde{e}_i^a)$ be total Nash-equilibrium emissions in the linked cap-and-trade system. Then the expression in the first term of the left hand side of Eq. (10) can be evaluated in a Nash equilibrium as

$$\left[\frac{\partial}{\partial e_i^c} \left\{ P(e_i^c + \sum_{j \neq i} \tilde{e}_j^c) \cdot (e_i^c - e_i^a) \right\} \right]_{e_i^c = \tilde{e}_i^c; e_i^a = \tilde{e}_i^a} = P(\tilde{E}) \cdot \left(1 - \frac{\partial e_i^a}{\partial e_i^c} \right) + P'(\tilde{E}) \cdot (\tilde{e}_i^c - \tilde{e}_i^a). \tag{11}$$

Next, substitute (11) into Eq. (10). After rearranging terms, we then derive

$$d_i + \left[P(\tilde{E}) + C'_i(\tilde{e}_i^a) \right] \frac{\partial e_i^a}{\partial e_i^c} - P(\tilde{E}) - P'(\tilde{E}) \cdot (\tilde{e}_i^c - \tilde{e}_i^a) = 0. \tag{12}$$

The term within the square brackets on the left hand side of (12) is zero because price equals marginal abatement cost in a cap-and-trade system. The first-order condition (12) then becomes

$$d_i - P(\tilde{E}) - P'(\tilde{E}) \cdot (\tilde{e}_i^c - \tilde{e}_i^a) = 0. \tag{13}$$

Add up over all i the expression (13), yielding the equation

$$\sum d_i - nP(\tilde{E}) - P'(\tilde{E}) \cdot \sum (\tilde{e}_i^c - \tilde{e}_i^a) = 0. \tag{14}$$

In equilibrium,

$$\sum \tilde{e}_i^a = \sum \tilde{e}_i^c, \tag{15}$$

so that the last term of the left hand side of Eq. (14) vanishes, turning Eq. (14) into

$$\sum d_i - nP(\tilde{E}) = 0. \tag{16}$$

Define \bar{d} to be the *average* marginal damage across all n jurisdictions

$$\bar{d} \equiv \frac{\sum d_i}{n}, \tag{17}$$

and let $\tilde{P} \equiv P(\tilde{E})$ be the equilibrium price of permits in the linked cap-and-trade system. Then we have from (16) and (17) the fundamental result that

$$\tilde{P} = \bar{d}. \tag{18}$$

It should be appreciated that Eq. (18) has been derived under extremely general assumptions about the abatement cost functions. The only substantive assumption, which accounts for the utter simplicity of expression (18), is the eminently defensible specification that marginal damages are constant for CO₂ emissions within a relatively short (5–10 year) period.

5 Linked versus Unlinked Cap-and-Trade

We have already derived for the unlinked voluntary cap-and-trade system that the self-volunteered autarchic price of permit within jurisdiction i is $\hat{p}_i = d_i$. Define \hat{P} to be the *average* unlinked voluntary permit price over all jurisdictions

$$\hat{P} \equiv \frac{\sum \hat{P}_i}{n} \tag{19}$$

Then, making use of (5), (17), and (19), Eq. (18) can be rewritten as

$$\tilde{P} = \hat{P}, \tag{20}$$

which means that the linked voluntary permit price is the *average* of the unlinked voluntary permit prices. Equations (18) and (19) provide a clear analytical image of the equilibrium price of permits in the linked cap-and-trade system.

We have repeatedly relied on the simplifying, but justified, assumption that, within a relatively short period (say 5–10 years), marginal damages are constant for each jurisdiction. Using this simplifying assumption again, how a linked jurisdiction sets its voluntary caps relative to actual emissions (and whether the jurisdiction buys or sells CO₂ permits) is fully characterized by a simple condition depending on the relationship between the jurisdiction’s marginal damages and the average marginal damages of the entire linked system.

From (18), $\tilde{P}(E) \equiv \tilde{P} = \bar{d}$, and then Eq. (13) can be rewritten as

$$\tilde{e}_i^c - \tilde{e}_i^a = k \cdot (\bar{d} - d_i), \tag{21}$$

where

$$k \equiv \frac{-1}{P'(\tilde{E})} > 0 \tag{22}$$

is viewed by all jurisdictions *i* as the same positive constant of proportionality.

Equation (21) is revealing. The only instrument under direct control of jurisdiction *i* is its voluntary cap e_i^c . Controlling the setting of its own voluntary cap e_i^c is the only way for jurisdiction *i* to influence total emissions \tilde{E} . In Nash equilibrium, $e_i^c = \tilde{e}_i^c$. It is then natural to ask: When does jurisdiction *i* set its control cap $e_i^c (= \tilde{e}_i^c)$ relatively low and when does jurisdiction *i* set its control cap $e_i^c (= \tilde{e}_i^c)$ relatively high? The natural benchmark for the voluntary setting of cap $e_i^c (= \tilde{e}_i^c)$ is a comparison with the actual post-cap-and-trade emissions of *i* – namely \tilde{e}_i^a . For all jurisdictions *i*, the actual equilibrium emissions \tilde{e}_i^a are a natural standard for comparison with equilibrium cap \tilde{e}_i^c because $\sum \tilde{e}_i^c = \sum \tilde{e}_i^a$ and because $\{\tilde{e}_i^a\}$ represents actual emissions normed to the *same* common price \tilde{P} . Equation (21) tells us exactly what is the sought-after difference ($\tilde{e}_i^c - \tilde{e}_i^a$).

From Eq. (21), we have the quantitative result that $\tilde{e}_i^c - \tilde{e}_i^a$ is directly proportional to $\bar{d} - d_i$ with the same constant of proportionality $k > 0$ for all *i*. This implies two qualitative results:

$$d_i > \bar{d} \iff \tilde{e}_i^c < \tilde{e}_i^a, \tag{23}$$

and

$$d_i < \bar{d} \iff \tilde{e}_i^c > \tilde{e}_i^a. \tag{24}$$

The interpretation of condition (23) should be relatively clear. When the marginal damage d_i to jurisdiction *i* is greater than the average marginal damages of the entire linked system \bar{d} , then jurisdiction *i* wants its voluntary cap $e_i^c = \tilde{e}_i^c$ to be relatively lower, in order to cut back total emissions even though, in this case, jurisdiction *i* ends up spending cash to buy

$(\tilde{e}_i^a - \tilde{e}_i^c) > 0$ emissions permits to counter-balance the relatively low setting of its voluntary cap.

The interpretation of condition (24) should likewise be relatively clear [but in the opposite direction of (23)]. When the marginal damage d_i to jurisdiction i is less than the average marginal damages of the entire linked system \bar{d} , then jurisdiction i wants its voluntary cap $e_i^c = \tilde{e}_i^c$ to be relatively higher, in order to allow total emissions to be higher, a direction toward which jurisdiction i is relatively tolerant because d_i is relatively low. In this case, jurisdiction i ends up with more cash by selling $(\tilde{e}_i^c - \tilde{e}_i^a) > 0$ emissions permits to counter-balance the relatively high setting of its voluntary cap, which is part of its motivation to issue relatively higher emissions permits.

How do total emissions compare between linked and unlinked voluntary cap-and-trade systems? Going back to Eq. (3), we can derive a simple condition for comparing total emissions, which, unfortunately, is not so simple to understand completely. A linked cap-and-trade system emits *less* in total than an unlinked cap-and-trade system if

$$\sum e_i(d_i) > \sum e_i(\bar{d}), \quad (25)$$

and conversely a linked voluntary cap-and-trade system emits *more* in total than an unlinked cap-and-trade system if

$$\sum e_i(d_i) < \sum e_i(\bar{d}). \quad (26)$$

Each of the n terms of (25) and (26) can be signed. If $d_i > \bar{d}$, then $e_i(d_i) < e_i(\bar{d})$ (with interpretation $\hat{e}_i < \tilde{e}_i^a$). Conversely, if $d_i < \bar{d}$, then $e_i(d_i) > e_i(\bar{d})$ (with interpretation $\hat{e}_i > \tilde{e}_i^a$). This signing of $e_i(\cdot)$ terms might be interpreted as hinting that the right and left hand sides of (25) and (26) might not differ greatly from each other since roughly half of the jurisdictions have the inequality in $e_i(\cdot)$ going one way and roughly the other half have the inequality in $e_i(\cdot)$ going the other way. However, this is merely crude heuristic hand-waving, not a formal argument.

Conditions (25) and (26) are not easy to analyze rigorously and could go either way, depending here on the distribution of the $\{d_i\}$ and the functions $\{e_i(\cdot)\}$. The literature is not decisive on this issue. Plausible arguments have been made on both sides.⁶

It is also difficult to characterize in general whether a jurisdiction has higher welfare from joining a linked cap-and-trade system or from remaining autarchic. It might have been presumed on basic principles of trade theory that joining a linked cap-and-trade system (offering a quasi-constant permit price of emissions) delivers higher welfare to a jurisdiction than remaining at the fixed autarchic level of emissions. However, this presumption does not hold in general for the situation here, where jurisdictions are gaming the system by strategically setting their own tradable emissions caps.⁷

⁶ For example, Holtmark and Sommervoll (2012) plausibly argue that total emissions are likely to be *higher* under linked voluntary cap-and-trade than under unlinked voluntary cap-and-trade. On the other hand, Carbone et al. (2009) plausibly argue that total emissions are likely to be *lower* under linked voluntary cap-and-trade than under unlinked voluntary cap-and-trade. Both of these two examples involve static games (as does this paper). In a dynamic model with both fossil fuels and renewables, Holtmark and Midtømme (2016) argue that linking leads to lower emissions. At this stage we think it is an open question how linking of cap-and-trade systems influences emissions.

⁷ Section III of Godal and Holtmark (2011) contains just such a counterexample where joining a linked cap-and-trade system yields lower welfare to a jurisdiction than remaining at the fixed autarchic level of emissions. In this paper we restrict ourselves to comparing linked and unlinked cap-and-trade systems without inquiring deeply into the individual motivations for participating in a linked system.

Table 1 Estimates of the jurisdictions' marginal damage costs. USD/tCO₂

Jurisdiction	Reference case	Damages based on RICE 2010	Average damages of three models
Brazil	3.1	2.4	2.9
Japan	4.8	2.3	2.4
EU	18.5	12.1	13.8
Canada	1.6	0.9	1.0
US	17.0	10.2	10.6
Russia	3.5	3.4	3.5
India	6.5	11.6	11.7
South Africa	0.7	0.7	0.7
China	14.8	15.8	11.0
Average	7.8	6.6	6.4

Source: Nordhaus (2015, the online appendix, Table B-2, p. 15)

6 Numerical Illustrations

In this section, the theoretical results of the previous sections are illustrated by reporting a set of model simulations in which nine jurisdictions link their cap-and-trade systems according to the set of rules that were assumed in the preceding sections.

6.1 Data Sources and Calibration

A partial equilibrium model was calibrated based on two data sources. While the parameters of the abatement cost functions and uncontrolled emissions are based on data provided by McKinsey (2009), estimated marginal damage costs from CO₂-emissions are based on Nordhaus (2015).

Nordhaus's model included six regions in addition to the EU, Brazil, Japan, Canada, the USA, Russia, India, South Africa and China. Because our theoretical approach is applicable to a set of independent jurisdictions, our model includes the eight mentioned countries and the EU, while the six (other) *regions* in Nordhaus's model were not included. The nine jurisdictions included in the model together represent approximately 72 percent of global CO₂-emissions.⁸

Table 1 shows the estimated marginal damage costs of the included jurisdictions. The first column includes estimates that assume that marginal damages are proportional to national GDPs. The second column shows estimates of marginal damages from the RICE 2010 model. The third column shows estimates of marginal damages based on the average of three widely used models (DICE-RICE, FUND, PAGE).

As in Nordhaus (2015), the set of GDP-based marginal damage cost estimates (the first column of Table 1) was used as the reference case. Results from simulations based on the two other sets of marginal damage costs are reported and discussed more briefly than the reference case.

The quadratic abatement cost functions are:

⁸ BP Statistical Review of World Energy 2018.

Table 2 Estimates of the regions' uncontrolled emissions (measured in GtCO₂/year), abatement potential at a cost of up to €60/tCO₂, and calibrated cost parameters

Jurisdiction	Uncontrolled emissions	Abatement potential	Parameter c_i
Brazil	3.1	1.9	34.6
Japan	1.5	0.3	219.2
EU	5.3	1.2	54.8
Canada	0.8	0.2	328.8
US	7.7	2.0	32.9
Russia	2.9	0.7	93.9
India	3.3	1.0	65.8
S.Africa	0.6	0.2	328.8
China	13.9	3.5	18.8

Source: Uncontrolled emissions and abatement potentials are taken from the 2020-case of McKinsey (2009). In the calibration of cost parameters, the euro-USD exchange rate was set at 1.096, as of medio September 2019

$$C_i(e_i) = \frac{c_i}{2}(e_i^u - e_i)^2, \quad (27)$$

where c_i is a parameter, e_i^u and e_i represent uncontrolled and actual emissions of jurisdiction i , respectively. Corresponding to (3), emissions without and with linking, respectively, become:

$$\hat{e}_i = e_i^u - \frac{d_i}{c_i}, \quad (28)$$

$$\bar{e}_i^a = e_i^u - \frac{\bar{d}}{c_i}. \quad (29)$$

It follows that if an emission level \bar{e}_i corresponds to marginal abatement cost \bar{p}_i , then the parameter c_i can be calibrated using the following formula:

$$c_i = \frac{\bar{p}_i}{e_i^u - \bar{e}_i}. \quad (30)$$

McKinsey (2009) included both a 2020 scenario and a 2030 scenario. Originally the model was calibrated to both the 2020 and the 2030 data and both cases were simulated. In the following, however, only the results from the model calibrated to the 2020 data are reported because results of the two exercises were quite similar.

The first column of Table 2 shows the estimates of the 9 jurisdictions' uncontrolled emissions in 2020. The second column shows estimated potential abatement that could be achieved if per-unit abatement costs do not exceed €60/tCO₂ in 2020.⁹ The last column of Table 2 shows the calibrated parameter values when the calibration formula given by Eq. (30) was applied to this data set.

⁹ McKinsey (2009, p. 148) states that "Following the IPCC definitions, the abatement cost curve shows technical measures with economic potential under 60 €/tCO₂e."

Table 3 Simulations results. Unlinked cap-and-trade systems. Reference case. Emissions in GtCO₂/year. Costs and welfare in billions USD/year

Jurisdiction	Emissions $e_i(d_i) = e_i^u - \frac{d_i}{c_i}$	Reductions of damage costs*	Abatement costs	Effects on welfare**
Brazil	3.01	5.9	0.139	5.7
Japan	1.48	9.1	0.053	9.1
EU	4.96	35.1	3.123	32.0
Canada	0.80	3.0	0.004	3.0
US	7.18	32.2	4.395	27.9
Russia	2.86	6.6	0.065	6.6
India	3.20	12.3	0.321	12.0
South Africa	0.60	1.3	0.001	1.3
China	13.11	28.1	5.829	22.2
Sum	37.20	133.7	13.929	119.8

* Reduction of damage costs compared to the case with uncontrolled emissions

** Increased welfare compared to the case with uncontrolled emissions

Table 4 Simulation results. Linked cap-and-trade systems. Reference case. Emissions and caps in GtCO₂/year. Costs, cash inflows and welfare in billions USD/year

Jurisdiction	Emissions $e_i(\bar{d}) = e_i^u - \frac{\bar{d}}{c_i}$	Reductions of damage costs*	Caps \tilde{e}_i^c	Net cash inflow $P \cdot (\tilde{e}_i^c - \tilde{e}_i^u)$	Abatement costs	Effects on welfare**
Brazil	2.87	4.1	3.67	6.2	0.886	9.4
Japan	1.46	6.3	1.97	4.0	0.140	10.1
EU	5.16	24.2	3.37	-14.0	0.560	9.7
Canada	0.78	2.1	1.82	8.2	0.093	10.2
US	7.46	22.3	5.93	-12.0	0.933	9.3
Russia	2.82	4.6	3.54	5.7	0.327	9.9
India	3.18	8.5	3.40	1.7	0.467	9.8
South Africa	0.58	0.9	1.77	9.3	0.093	10.2
China	13.48	19.4	12.32	-9.1	1.633	8.6
Sum	37.79	92.4	37.79	0.0	5.132	87.2

* Reduction of damage costs compared to the case with uncontrolled emissions

**Increased welfare compared to the case with uncontrolled emissions

6.2 Simulation Results

In the unlinked case, jurisdiction *i* seeks to *minimize* over e_i the expression $D_i(E) + C_i(e_i)$, which gives emissions \hat{e}_i . Results of simulation of the unlinked reference case are given by Table 3.

In the linked case, jurisdiction *i* seeks to *maximize* over e_i^c the expression $P(E) \cdot (e_i^c - e_i^u) - (D_i(E) + C_i(e_i^u))$ and the solution is denoted \tilde{e}_i^c . With the quadratic abatement cost function specified in (27), the first-order condition (13) gives the following voluntary caps with linking:

Table 5 Total emissions and welfare in the cases with linked voluntary cap-and-trade systems. Change from the case without linking

Effects on:	Reference case	Case with damages according to RICE 2010	Case with average damages of three models
Emissions (percentage change)	+1.6	+1.5	+1.1
Welfare (billion USD/year)	-32.56	-26.48	-18.80

$$\tilde{e}_i^c = \hat{e}_i - \sum_{j \neq i} \frac{1}{c_j} (d_i - \bar{d}). \quad (31)$$

The results of simulating the reference case with linking are given in Table 4. The average marginal damage cost in the reference case is 7.8 USD/tCO₂. This marginal cost becomes the equilibrium price of permits in the linked cap-and-trade system, see (18).

From Eq. (21) it follows that the number of permits sold by jurisdiction i , $\tilde{e}_i^c - \tilde{e}_i^a$, is directly proportional to $\bar{d} - d_i$ with the same constant of proportionality k for all i . Given the numerical assumptions, $k = 0.167$.

According to expression (23), jurisdictions with marginal damage costs, d_i , higher than average, set their voluntary cap lower than their total emissions and end up buying permits. This applies to the EU, the US and China, in the reference case. These jurisdictions end up with negative cash inflows, see Table 4. Because the permit price with linking is lower than the marginal damage costs of these three jurisdictions, their emissions are higher with linked cap-and-trade systems compared to the unlinked case. At the same time, their caps are set tighter. The reason for this is that the cost of reducing (global) emissions on the margin is reduced for these jurisdictions when they link their markets with the other jurisdictions.

Six jurisdictions have in the reference case marginal damage costs lower than average, and opposite mechanisms come into play. In accordance with (24), they end up selling permits and collect positive cash inflows. The emissions of these six jurisdictions are lower with linking compared to the unlinked case. Note also that their caps are even greater than their uncontrolled emissions, see Tables 2 and 4.¹⁰

In the reference case, the total emissions of the nine jurisdictions are 1.6 percent higher with linked cap-and-trade systems compared to the cases without linking, see Table 5. This table also shows the effects on total emissions when the two other sets of damage cost estimates are applied. In all cases, linking leads to increased emissions. However, as indicated in our discussion in relation to the conditions (25) and (26), the total emissions for linked voluntary cap-and-trade systems do not differ very much from the emissions for unlinked systems.

Linking provides an efficient allocation of abatement efforts, with lower total abatement costs compared to the unlinked case. At the same time, linking results in higher total emissions. Because total emissions already in the first place were inefficiently high, this represents a welfare loss. It turns out that the welfare losses due to increased emissions with linking in all three cases considered are greater than the reduction of abatement costs from efficient allocation of abatement efforts. In all cases considered, welfare is therefore lower with linking compared to the unlinked cases; see the bottom line of Table 5.

¹⁰ Helm (2003) pointed out that this could be the result.

7 Concluding Remarks

An appropriate and widely accepted specification for the marginal damages of CO₂ emissions within a relatively short (5–10 years, say) period is that they are constant for each jurisdiction. This critical, but defensible, assumption greatly clarifies the analysis of linkage of cap-and-trade systems and yields simple closed-form expressions for all (linked and unlinked) CO₂ emissions prices. The current paper has derived and discussed some implications of this simplicity for linked and unlinked voluntary CO₂ cap-and-trade systems. How a linked jurisdiction sets its voluntary caps relative to actual emissions (and whether the jurisdiction buys or sells CO₂ permits) is fully characterized by a simple linear proportionality condition that depends only on the difference between the jurisdiction's marginal damages and the average marginal damages of the entire linked system. Whether linkage increases or decreases overall emissions depends on a condition that is easy to express but difficult to evaluate rigorously, and the answer could go either way.

Because our theoretical results were ambiguous, we calibrated and simulated a partial equilibrium model which includes nine jurisdictions that in 2017 together were responsible for 72.8 percent of global CO₂-emissions (BP Statistical Review of World Energy 2018). We applied three different sets of marginal damage costs provided by Nordhaus (2015). In all three considered cases, the simulations of linkage resulted in slightly greater overall emissions compared to the simulations of unlinked national cap-and-trade markets.

Acknowledgements Open Access funding provided by Statistics Norway.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Bohm P (1992) Distributional implications of allowing international trade in CO₂ emission quotas. *World Econ* 15(1):107–114
- Carbone JC, Helm C, Rutherford TF (2009) The case for international emission trade in the absence of cooperative climate policy. *J Environ Econ Manag* 58(3):266–280
- Chen H (2017) The Paris agreement on climate change (Issue Brief November 2017). Natural Resources Defense Council. <https://www.nrdc.org/resources/paris-agreement-climate-change>
- Eyckmans J, Proost S (1996) Voluntary provision of greenhouse gas abatement and emission trading equilibria. In: Gerland, Gorka (eds) *The economics of atmospheric pollution*. Springer, Berlin, pp 69–98
- Godal O, Holtsmark B (2011) Permit trading: merely an efficiency-neutral redistribution away from climate-change victims? *Scand J Econ* 113(4):784–797
- Green JF, Sterner T, Wagner G (2014) A balance of bottom-up and top-down in linking climate policies. *Nat Clim Change* 4:1064
- Helm C (2003) International emissions trading with endogenous allowance choices. *J Public Econ* 87(12):2737–2747
- Holtsmark B, Sommervoll DE (2012) International emissions trading: good or bad? *Econ Lett* 117(1):362–364
- Holtsmark K, Midttømme K (2016) The dynamics of linking permit markets (Working Paper). University of Oslo, Department of Economics. https://sv.uio.no/econ/english/people/aca/katinkah/permits_holtsmark_midttomme_july16.pdf

- Krishna K, Tan LH (1999) Transferable licenses versus nontransferable licenses: what is the difference? *Int Econ Rev* 40(3):785–800
- Levin K, Fransen T (2015) With latest climate commitments, how much will the world warm? It's complicated (Web article). World Resources Institute. <https://www.wri.org/blog/2015/11/latest-climate-commitments-how-much-will-world-warm-its-complicated>
- McKinsey (2009) Pathways to a low-carbon economy. Version 2 of the global greenhouse gas abatement cost curve. www.mckinsey.com (Report)
- Mehling MA, Metcalf GE, Stavins RN (2017) Linking heterogeneous climate policies (Consistent with the Paris Agreement). Harvard Project on Climate Agreements
- Nordhaus W (2015) Climate clubs: overcoming free-riding in international climate policy. *Am Econ Rev* 105(4):1339–70
- UNFCCC (2015) The Paris Agreement (COP21). https://unfccc.int/sites/default/files/english_paris_agreement.pdf

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.