Designing Emissions Trading Schemes: Negotiations on the Amount & Allocation of Permits*

Tom Rauber† Fabian Naumann‡

August 27, 2024

Abstract

National free-riding incentives prevent necessary reductions in global emissions and thus paralyze combating climate change. By considering two countries that engage in bargaining over a joint emissions cap according to Rubinstein's (1982) alternating-offers model, we discover a simple mechanism to overcome free-riding incentives and achieve efficient outcomes. Provided that countries are sufficiently symmetric, allowing them to endogenously design a joint cap-and-trade system by negotiating the amount and allocation of permits yields the efficient level of emissions and maximizes welfare. In contrast, if the negotiating countries are too heterogeneous, the scope for implicit side payments in this system may not be sufficient to implement an efficient emissions cap.

Keywords: Emissions Trading Schemes, Alternating-Offers Model, Nash Bargaining Solution, Outside Option, Free-Riding.

JEL Classification: C78, H41, Q53, Q58.

Declarations of Interest: None.

*Acknowledgments. We thank Joshua Bißbort, Daniel Heyen, Çağıl Koçyiğit, Paul Ritschel, Robert Schmidt, Joseph Shapiro, Philipp Weinschenk, and Jan Wenzelburger, as well as the participants of research seminars at the Universities of Cambridge, Kaiserslautern-Landau, and Luxembourg, for helpful comments and suggestions.

†Faculty of Business Studies & Economics, University of Kaiserslautern-Landau, Gottlieb-Daimler-Straße 42, D-67663 Kaiserslautern, Germany, tom.rauber@wiwi.uni-kl.de (corresponding author).

‡Faculty of Business Studies & Economics, University of Kaiserslautern-Landau, Gottlieb-Daimler-Straße 42, D-67663 Kaiserslautern, Germany, fabian.naumann@wiwi.uni-kl.de.

1 Introduction

Climate change constitutes one of the most severe challenges currently facing humanity (see, among many others, Stern, 2007; Weitzman, 2007; Nordhaus, 2019). It is well-known that extensive global emissions of greenhouse gases such as carbon dioxide, methane, and nitrous oxide lead to an increase in the temperature at the Earth's surface, which in turn promotes weather and climate extremes worldwide. Some future changes in the climate system are already inescapable, but their extent can be limited via immediate and effective emissions cuts (IPCC, 2023). While lower global emissions are thus undoubtedly beneficial from a normative perspective[1], the incentive to free-ride on other countries' abatement activities has created a deadlock (Underdal, Hovi, Kallbekken, & Skodvin, 2012). An emissions trading scheme ("cap-and-trade system") is a prominent policy instrument that, if designed appropriately, breaks the deadlock and yields efficient outcomes (Schmalensee & Stavins, 2017). However, determining, implementing, and enforcing an appropriate design is a major challenge for policy-makers (Egenhofer, 2007; Nordhaus, 2007; Stavins, 2008a; Weitzman, 2014). Our paper explores whether a simple mechanism can implement efficient outcomes: Allowing countries to endogenously design a joint cap-and-trade system through negotiations. We thus shed light on the question of whether this simple procedure for designing emissions trading schemes can solve the problem of excessive greenhouse gas emissions.

We consider a simple two-country model in which each country bears individual costs for reducing emissions, while, at the same time, benefiting not only from its own abatement but also from abatement activities carried out in the other country. The countries are rational and negotiate according to the alternating offers a la Rubinstein (1982) with complete and perfect information. Due to the negligible friction in the bargaining process, the subgame perfect equilibrium in this dynamic game coincides with Nash's (1950) bargaining solution, i.e., the solution to a simple optimization problem.

Our analysis starts by exploring the benchmark scenario in which both countries deploy national caps. Due to the positive externality – countries benefit from each other's reduction in emissions – we find a strong free-riding incentive that leads to an inefficiently high level of emissions. We then direct attention to a setting

[1] Stern (2008, p. 1) even refers to the excessive greenhouse gas emissions as *"the biggest market failure the world has seen."*

where the countries set a joint emissions cap through negotiations. The basic mechanisms are explored in a stylized setting where the countries have agreed on setting up a joint cap-and-system and negotiate only on the amount of certificates. To derive our main results, we then successively add more degrees of freedom by *(i)* allowing the countries not only to bargain on the amount of certificates but also on their initial allocation. Afterwards, we *(ii)* incorporate the possibility of strategic termination of the negotiations. That is, each country may actively decide to end the negotiations, resulting in national abatement activities. A comparison with the benchmark then yields our two main results.

First, enabling the countries to set up a joint emissions trading scheme and to bargain over its design allows them to overcome the free-riding incentive and implement the efficient emissions cap if the countries are sufficiently symmetric. That is, if they are sufficiently similar in terms of their cost and benefit structures as well as their initial emissions. This holds irrespective of the possibility of strategic opting out of the negotiations. Intuitively, the countries agree on the cap that maximizes overall welfare. The initial allocation of certificates is then used as implicit side payment to distribute the highest level of welfare equally across countries. Second, negotiations do not necessarily implement the efficient cap if countries' benefits resulting from this cap are too heterogeneous. The intuition is as follows. The scope for using the initial allocation as implicit side payment is limited as a country cannot receive more than the entire share of certificates. If the benefits from the efficient cap are sufficiently different, then even allocating all certificates to the low-benefit country is insufficient to generate equal welfare levels in both countries. The countries will then rather agree on another cap that is less efficient but yields a more equal distribution of individual welfare levels. As we argue afterwards, our two insights also carry over to settings in which more than two countries negotiate on the design of a multilateral cap-and-trade system.

Related Literature. Free-riding incentives have received particular attention in the literature as a primary force preventing emissions reduction and hindering strict international agreements to curb climate change (see, e.g., Carraro & Siniscalco, 1993; Barrett, 1994, 2003; Nordhaus, 2015). They result from the public good nature of abatement. While all countries benefit from lower global emissions, only those countries that actually reduce their emissions bear the associated costs. This situation gives rise to Hardin's (1968) infamous tragedy of commons: Each country will leave costly abatement activities to the others. As a remedy for the free-riding problem, economic theory proposes *price-based* and *quantity-based* instruments (see, e.g., Weitzman, 1974; Mas-Colell, Whinston, & Green, 1995; Nordhaus, 2007).[2] While price-based instruments originate from Pigou's (1920) taxation of externalities, quantity-based instruments build on the concept of tradable permits as proposed by Coase (1960) and Dales (1968). Nowadays, both approaches are widespread policy instruments to address either national or international climate targets.

For price-based instruments, Weitzman (2014) explores a simple mechanism to alleviate the free-riding incentive. Given that the countries can commit to a single emissions price, they determine this price by pairwise majority voting, resulting in emissions close to the efficient level. Intuitively, due to global commitment to the emissions price, a country's additional costs from a higher emissions price are offset by its additional benefit arising since all other countries reduce their emissions at the same time in response to the higher price. However, for quantity-based instruments, Weitzman (2014, p. 31) suspects that *"even if there were a collective commitment to negotiate or vote on a second-stage worldwide total emissions cap, disagreements over the first-stage subdivision formula (...) would paralyze such a quantity-based approach."* Despite these doubts, our analysis reveals that a similar mechanism also exists for quantity-based instruments where the total amount of emissions is capped, and the emissions price is determined via trading, given that the countries are sufficiently symmetric. We thereby contribute to two strands of the emissions-trading literature, namely endogenous allowance choices and linking emissions trading schemes.

Previous research by Helm (2003) has shown that *endogenous allowance choices* made by countries do not automatically result in lower pollution levels, as environmentally more (less) concerned countries choose to pollute less (more) such that the environmental efforts offset. Using numerical simulations, Smead, Sandler, Forbes, and Basl (2014) investigate a setup where agents bargain over their share of the fixed total emissions. They find that negotiations tend to fail if too many agents request over-proportional emissions shares, making the initial demand for those shares a key factor for successful negotiations. Our paper complements these results by identifying an insufficient scope for side payments as an additional mechanism that may prevent efficient endogenous allowance choices.

[2]See Aldy, Barrett, and Stavins (2003) for a more nuanced distinction and Goulder and Schein (2013) as well as Stavins (2022) for a comparison of the different approaches.

Considering situations where emissions trading systems have already been implemented, the literature on *linking emissions trading schemes* raises the question as to whether combining these systems is beneficial. Flachsland, Marschinski, and Edenhofer (2009) analyze the benefits and drawbacks of linking, such as reduced volatility, strengthening the multilateral commitment versus expanded emission caps, abatement targets that are not in line with a burden-sharing approach, and declining national regulatory power. Doda and Taschini (2017) argue that linking becomes more advantageous the larger the jurisdictions' size and variances of benefit shocks, while a stronger correlation of these shocks and higher sunk costs of linking result in the opposite effect. Doda, Quemin, and Taschini (2019) find that multilateral linking can lead to tremendous efficiency gains, which arise equally from effort- and risk-sharing. However, Habla and Winkler (2018) demonstrate that strategic delegation hinders the linking of emissions trading schemes.[3] We contribute to this literature by showing that linking arises naturally via negotiations if the countries are sufficiently symmetric.

More generally, applying game theory to analyze negotiations on ecological agreements and the provision of environmental public goods has been a vibrant research area over the past two decades (Caparrós, 2016).[4] Bargaining models were deployed in the context of global north-south climate change negotiations (cf. Caparrós, Péreau, & Tazdaït, 2004), air pollution (cf. Harstad, 2007), investment in green technologies (cf. Urpelainen, 2012), global biodiversity regulation (cf. Swanson & Groom, 2012), interregional water sharing (cf. Nehra & Caplan, 2022), and climate policies (cf. Harstad, 2023). The paper closest in spirit to ours is Dijkstra and Nentjes (2020), who analyze negotiations on tradable production certificates in a related model. While the structure of their game is different, their results also differ in so far as they find that bargaining *always* leads to efficient production levels.[5] In spite of this plentiful literature, this paper is, to the best of our knowledge, the first to explore how two countries endogenously design a joint cap-and-trade system by bargaining over the amount and allocation of permits.

[3]A strand in the political economy literature examines how delegates in the context of international environmental agreements are chosen. The baseline shared here is that countries may select delegates that misrepresent their preferences (see, e.g.,Segendorff, 1998; Graziosi, 2009; Habla & Winkler, 2018; Arvaniti & Habla, 2021). We, however, abstract from such considerations and assume that the countries' preferences are correctly represented in the negotiations.

[4]For an excellent overview of the literature, see Caparrós (2016).

[5]The exact same applies to Arvaniti and Habla (2021) in the political economy literature.

Outline. The remainder of this paper is organized as follows. In the next section, we introduce our two-country model, derive basic insights about the abatement levels as well as the endogenous emissions market price, and define a country's welfare as a function of the design of the emissions trading scheme. Section 3 analyzes two benchmark scenarios for our welfare analysis by considering a social planner and national caps. In Section 4, we explore the simplest version of cap negotiation, i.e., a setting where the countries can bargain on the emissions cap only. Section 5 then generalizes the preceding analysis by allowing the countries to bargain simultaneously on the emissions cap and the initial allocations of permits, while also incorporating the possibility of strategic opting out of the negotiations. A discussion of multilateral negotiations, i.e., the case of more than two countries, is provided in Section 6. Finally, Section 7 concludes our analysis by highlighting its implications. All proofs are provided in the Appendix.

2 Model & Basic Insights

This section presents the underlying theoretical framework for describing a cap-and-trade system. Subsequently, we derive individual welfare for a country as a function of the scheme's design.

Framework. Consider two countries that emit greenhouse gases. Each country $i \in \{1, 2\}$ carries out abatement activities a_i relative to its emissions level under "business as usual" $e_{i,0} \in \mathbb{R}_{++}$.[6] Abatement activities affect country i's welfare through three channels. First, there are benefits of overall abatement $B_i(\sum_i a_i)$. Reducing emissions has thus a positive externality since country i also benefits from the abatement made by country $-i$ and vice versa. Second, country i bears costs of its own abatement $C_i(a_i)$. Third, as countries participate in a cap-and-trade system, emissions trading additionally results in either revenues or costs, depending on whether i is a seller or buyer of permits. A country acts as a seller [buyer] of permits if its realized emissions e_i are lower [higher] than its initial endowment of permits $\bar{e}_i \in \mathbb{R}_+$. The emissions market price p is endogenously

[6] As we will see later, our assumptions ensure that $a_i \geq 0$, i.e., a country does not increase its emissions above the initial level generated under "business as usual".

determined. Putting these components together, country i's welfare amounts to:

$$B_i(\sum_i a_i) - C_i(a_i) + (\bar{e}_i - e_i) \cdot p, \qquad i = 1, 2. \quad (2.1)$$

Note that the benefits of lower emissions are expressed in terms of emissions abatement, i.e., the higher the abatement, the lower the emissions and thus the higher the benefits. For the emissions cap of the entire scheme, $\bar{E}$, it holds that

$$\bar{E} = \sum_i \bar{e}_i \qquad \text{and} \qquad \bar{e}_i = \mu_i \bar{E} \qquad i = 1, 2, \quad (2.2)$$
$$\text{where} \quad \mu_2 = 1 - \mu_1,$$

and $\mu_1 \in [0, 1]$ denoting the share of permits allocated to country 1. As realized emissions are determined by emissions under "business as usual" minus abatement, we can rewrite the country i's welfare in (2.1) as

$$B_i(\sum_i a_i) - C_i(a_i) + (\mu_i \bar{E} - (e_{i,0} - a_i))p, \qquad i = 1, 2. \quad (2.3)$$

Following the literature, in which abatement costs are commonly assumed to be quadratic (see, for instance, Weitzman, 1974; Barrett, 1994; McGinty, 2007; Weitzman, 2014; Gersbach & Hummel, 2016; Baudry, Faure, & Quemin, 2021), we impose the following assumptions on the abatement costs:

Assumption 1 (Cost Function).
Country i's abatement costs are captured by a cost function $C_i : [0, e_{i,0}] \rightarrow \mathbb{R}_+$, which is of the quadratic form

$$C_i(a_i) = \frac{\zeta_i}{2} a_i^2, \quad \textit{where} \quad \zeta_i > 0, \qquad i = 1, 2. \quad (2.4)$$

Without loss of generality, the indices are such that country 1 has weakly higher marginal abatement costs, i.e., $\zeta_1 \geq \zeta_2$. We further normalize that $\zeta_1 + \zeta_2 = 1$ to simplify the exposition. By defining $A := \sum_i a_i$ and $E_0 := \sum_i e_{i,0}$, country i's benefit writes as $B_i(A)$ and satisfies the following assumptions.

Assumption 2 (Benefit Function).
Country i's benefit function $B_i : [0, E_0] \rightarrow \mathbb{R}_+$ is twice continuously differentiable, strictly increasing, strictly concave, and satisfies

$$B_i(0) = 0, \quad B_i'(0) = \infty, \quad B_i'(E_0) = 0, \quad \textit{and} \quad B_i'' < -\zeta_i \zeta_{-i}^2, \qquad i = 1, 2.$$

The assumptions on the marginal benefits ensure that reducing emissions relative to "business as usual" is beneficial for both countries. Moreover, the assumption on the second derivative is a technical one, guaranteeing that country i's welfare function, which will be introduced shortly, is concave and thus well-behaved.

Abatements & Welfare Functions. We start by deriving fundamental insights about the realized emissions market price and the abatement activities within a given cap-and-trade system. As we consider a unit-mass continuum of homogeneous price-taking firms in each country, a firm's problem consists of minimizing its costs under the emissions trading scheme. The optimization problem of a representative firm in country i is thus

$$\min_{0 \leq a_i \leq e_{i,0}} p \cdot (e_{i,0} - a_i) + C_i(a_i), \qquad i = 1, 2.$$

From the corresponding first-order condition (FOC), we directly obtain the optimal abatement level:

$$-p + C_i'(a_i) = 0 \quad \Longleftrightarrow \quad a_i = \frac{p}{\zeta_i}, \qquad i = 1, 2. \quad (2.5)$$

Since overall emissions are restricted by the emissions cap $\bar{E}$, market clearing in the emissions permit market requires that $\sum_i (e_{i,0} - a_i) = \bar{E}$. Inserting the abatements per country and solving for the emissions market price yields:

$$p = p(\bar{E}) := \zeta_1 \zeta_2 \left(E_0 - \bar{E} \right). \quad (2.6)$$

Plugging (2.6) into (2.5), we obtain the abatement activities and calculate their derivatives with respect to $\bar{E}$:

$$a_i = a_i(\bar{E}) := \zeta_{-i} \left(E_0 - \bar{E} \right), \quad \text{and} \quad \frac{\partial a_i(\bar{E})}{\partial \bar{E}} = -\zeta_{-i}, \qquad i = 1, 2, \quad (2.7)$$

$$A = A(\bar{E}) := E_0 - \bar{E}, \quad \text{and} \quad \frac{\partial A(\bar{E})}{\partial \bar{E}} = -1. \quad (2.8)$$

Since country 2 represents the country with lower abatement costs ($\zeta_2 \leq \zeta_1$), it contributes more to total abatement ($a_2 \geq a_1$). This is intuitive and in accordance with economic insights on emissions trading. By inserting (2.6)–(2.8) in expression (2.3), we are now in the position to define country i's welfare as a function of the design of the cap-and-trade system, i.e., the amount and allocation of permits:

$$W_i(\bar{E}, \mu_1) := B_i(A(\bar{E})) - C_i(a_i(\bar{E})) + (\bar{E}\mu_i - e_{i,0} + a_i(\bar{E}))p(\bar{E}), \quad i = 1, 2, \quad (2.9)$$

$$\text{where} \quad \mu_2(\mu_1) := 1 - \mu_1.$$

3 Benchmarks

We next examine two benchmark scenarios, against which we evaluate the endogenous design of the cap-and-trade system through negotiations. First, we consider a social planner who designs a joint emissions cap. Second, we investigate how each country would set its national cap individually if there was no joint cap-and-trade system. By comparing both of these scenarios, we then identify the welfare-reducing effect resulting from free-riding in our model.

Social Planner. Let us first consider a social planner who seeks to maximize the overall welfare of both countries. It is apparent that, from a social planner's perspective, trading activities between the countries offset each other, which renders the allocation μ_1 irrelevant for overall welfare. Hence, the optimization problem faced by the social planner is solely to choose a cap that maximizes welfare. Formally, the efficient cap $\bar{E}^S$ is defined by the solution to:

$$\max_{0\leq\bar{E}\leq E_0} W(\bar{E}) := \sum\nolimits_i W_i(\bar{E}, \mu_1). \tag{3.1}$$

The FOC to the social planner's problem is

$$\frac{dW(\bar{E})}{d\bar{E}} = \sum\nolimits_i \frac{dW_i(\bar{E}, \mu_1)}{d\bar{E}} = 0. \tag{3.2}$$

By inserting (2.9) and using the deviates in (2.7)–(2.8), Equation (3.2) can be simplified to

$$\sum\nolimits_i B_i'(A(\bar{E})) = C_i'(a_i(\bar{E})), \qquad i = 1, 2. \tag{3.3}$$

Indeed, solely balancing marginal cost of abatement with the overall marginal benefit of abatement is what determines the efficient cap. The following lemma establishes the existence and uniqueness of the efficient cap $\bar{E}^S$.

Lemma 1. *There exists a unique efficient cap* $0 < \bar{E}^S < E_0$ *that solves Problem* (3.1)*. It is determined by the solution to* (3.3)*.*

Intuitively, both countries benefit from a marginal increase in the abatement irrespective of where the emissions have been reduced (see l.h.s. of (3.3)). Since permits are traded, marginal abatement costs of the countries equalize (see r.h.s. of (3.3)) such that any emissions target is met at the lowest cost, making a cap-and-trade system an efficient policy instrument to regulate pollution.[7] Welfare is then maximized at the efficient emissions level $\bar{E}^S$ that balances the total marginal benefits of abatement with its marginal costs. We define the maximum level of welfare generated by the efficient cap as

$$W^S := W(\bar{E}^S).$$

National Caps. Next, we turn our attention to a scenario in which both countries do *not* participate in a joint emissions trading scheme but deploy national regulations in the form of national emissions caps instead. Since each country implements its own emissions cap, the corresponding abatement for country i and the corresponding overall abatement are of the form

$$a_i(\bar{e}_i) := e_{i,0} - \bar{e}_i, \qquad i = 1, 2, \quad (3.4)$$

$$A(\bar{e}_i, \bar{e}_{-i}) := \sum_i a_i(\bar{e}_i) = E_0 - \sum_i \bar{e}_i. \quad (3.5)$$

Given the cap of the other country, country i now chooses its own cap to maximize its welfare, i.e., by solving

$$\max_{0 \le \bar{e}_i \le e_{i,0}} B_i(A(\bar{e}_i, \bar{e}_{-i})) - C_i(a_i(\bar{e}_i)), \qquad i = 1, 2. \quad (3.6)$$

Using (3.4) and (3.5), country i's FOC can be written as

$$B_i'(A(\bar{e}_i, \bar{e}_{-i})) = C_i'(a_i(\bar{e}_i)), \qquad i = 1, 2. \quad (3.7)$$

Lemma 2 now establishes the existence and uniqueness of a Nash equilibrium in this abatement game.

Lemma 2. *For each country $i \in \{1, 2\}$, there exists a unique Nash equilibrium cap $0 \le \bar{e}_i^N < e_{i,0}$ that solves Problem (3.6). The first inequality is strict for at least one country.*

Intuitively, in the Nash equilibrium, the cap $\bar{e}_i^N$ chosen country i is the best response – by equating marginal benefits and costs – to the cap $\bar{e}_{-i}^N$ implemented by the other country. Hence, no country has an incentive to deviate. The overall welfare generated by national caps is defined by

$$W^N := \sum_i W_i^N, \quad \text{where} \quad W_i^N := B_i(A(\bar{e}_i^N, \bar{e}_{-i}^N)) - C_i(a_i(\bar{e}_i^N)), \qquad i = 1, 2.$$

[7] This holds true in the absence of transaction costs and imperfect competition (see, e.g., Hahn, 1984; Stavins, 1995).

Comparison. Naturally, the following question arises: How effective are national caps in reducing emissions and improving welfare in comparison to the efficient outcome generated by the social planner? The following proposition answers this question by comparing both benchmark scenarios in terms of implemented overall cap and welfare.

Proposition 1 (National Caps Versus Social Planner).
Compared to a social planner, national caps implement a strictly higher overall cap, $\sum_i \bar{e}_i^N > \bar{E}^S$*, that leads to a strictly lower level of welfare,* $W^N < W^S$*.*

Proposition 1 is the manifestation of the free-riding problem in our model. Intuitively, each country i has an incentive to free-ride on the abatement carried out by the other country, while, at the same time, implementing insufficient domestic abatement targets to reduce its own abatement costs. This results in total emissions that are too high from a societal perspective. Example 1 illustrates this finding.

Example 1 (Free-Riding).
Consider benefits that are captured by a benefit function of the form

$$B_i(A) = \beta_i(2\sqrt{A} - A), \quad \textit{where} \quad \beta_i > 2\zeta_i\zeta_{-i}^2, \quad \textit{and} \quad A \in [0, 1].$$

Even for the symmetric case where $e_{1,0} = e_{2,0} = 0.5$*,* $\zeta_1 = \zeta_2 = 0.5$*, and* $\beta_1 = \beta_2 = 0.3$*, free-riding leads to emissions that are 35% above the efficient emissions cap.*

4 Negotiations on the Emissions Cap

Our analysis of the benchmarks raises the question of whether a cap-and-trade system can still overcome or at least mitigate the free-riding incentive in the absence of a social planner. In other words, is it more favorable from a societal perspective if the countries design the cap-and-trade system themselves rather than implementing national caps?

Probably the most natural way for countries to endogenously determine the design of the cap-and-trade system is through negotiations, which we will analyze next. To understand the basic mechanisms, it is illustrative to start with the simplest form of negotiations, where the countries only bargain over the emissions cap. That is, we consider a setting where the countries have already *(i)* agreed on setting up a joint cap-and-trade system and *(ii)* determined a certain initial allocation of permits $\hat{\mu}_1 \in [0, 1]$.

(i) means that neither country can strategically opt out of the negotiation, which, for our bargaining model, implies that no country has an outside option. One rationale for this situation could be that governments have already committed to establishing a joint cap-and-trade system or that public pressure is forcing them to do so.

(ii) entails that the allocation of permits is exogenous from the perspective of the negations on the emissions cap. This allocation could, for instance, be determined through prior negotiations or *grandfathering*, i.e., proportional to the countries' emissions under "business as usual" going back to the concept of first possession and appropriation (cf. Epstein, 1979; Lueck, 1995; Rose, 1985).[8] As grandfathering is frequently practiced in emissions trading systems such as the US sulphur dioxide emissions trading program and the EU ETS (Woerdman, Arcuri, & Clò, 2008), the assumption of an exogenous allocation is plausible from a practical point of view.

Nonetheless, it should be emphasized that *(i)* and *(ii)* will be dropped in the course of this paper in order to derive more general insights.

4.1 Bargaining Outcome

First, we need to specify the set of feasible bargaining solutions. It is straightforward that any bargaining has to result in a Pareto-efficient outcome. Otherwise, the parties could simply agree on another cap and thereby achieve a Pareto improvement. To construct the set of Pareto-efficient caps, $\mathscr{P}_{\hat{\mu}_1}$, it is necessary to determine which joint cap $\bar{E}_i$ country i prefers most as an outcome of the bargaining procedure. Country i would set a global cap that maximizes its welfare,

$$\max_{0 \leq \bar{E} \leq E_0} W_i(\bar{E}, \hat{\mu}_1), \qquad i = 1, 2. \quad (4.1)$$

[8] Analyses of grandfathering from an economic perspective can be found, for example, in Böhringer and Lange (2005), Damon, Cole, Ostrom, and Sterner (2019), as well as Grimm and Ilieva (2013).

From differentiating (2.9) and simplifying, we obtain country i's FOC, which reads

$$- B_i'(A(\bar{E})) + \mu_i C_i'(a_i(\bar{E})) + x_{i,-i}(\bar{E})\ C_i''(a_i(\bar{E}))\frac{\partial a_i(\bar{E})}{\partial \bar{E}} = 0, \qquad i = 1, 2, \quad (4.2)$$

$$\text{where} \quad x_{i,-i}(\bar{E}) \coloneqq \left(\mu_i \bar{E} - e_{i,0} + a_i(\bar{E})\right).$$

As (4.2) shows, from an individual perspective, trading activities *and* the initial distribution of permits matter for welfare. Indeed, $x_{i,-i}$ is the amount of permits passed from country i to $-i$, which can be both positive and negative depending on whether i sells or purchases permits from country $-i$. The l.h.s. of Equation (4.2) reveals three marginal effects that a higher cap has on country i's welfare. There are effects on marginal benefits (first summand) and marginal costs (second summand), as well as a trading effect (third summand). The following lemma now establishes the existence and uniqueness of a solution to Problem (4.1) and compares it to the social planner solution.

Lemma 3. *For each country $i \in \{1, 2\}$, there exists a unique individually optimal cap $0 \leq \bar{E}_i < E_0$ that solves Problem* (4.1). *The first inequality is strict for at least one country. Moreover, it holds that min $\{\bar{E}_1,\ \bar{E}_2\} \leq \bar{E}^S \leq$ max $\{\bar{E}_1,\ \bar{E}_2\}$.*

Lemma 3 reveals that at most one country advocates complete decarbonization, while the other country prefers a positive level of global emissions. Positive caps are determined by the FOC (4.2) to balance marginal benefits, costs, and trading effects. A social planner, in comparison, would cap overall emissions at a level that lies between those caps optimal from an individual perspective. Since both countries' welfare functions are strictly concave in the implemented cap, we can directly use Lemma 3 to define the Pareto set $\mathscr{P}_{\hat{\mu}_1}$.[9]

Definition 1 (Pareto-Efficient Caps)**.**
The set of Pareto-efficient caps, $\mathscr{P}_{\hat{\mu}_1} \subset \mathbb{R}_+$, is defined by

$$\mathscr{P}_{\hat{\mu}_1} \coloneqq \left\{\bar{E} : \bar{E} \in \Big[\min\{\bar{E}_1,\ \bar{E}_2\},\ \max\{\bar{E}_1,\ \bar{E}_2\}\Big]\right\}.$$

Figure 1 illustrates the idea of Definition 1 for the case where $\bar{E}_1 < \bar{E}_2$. Caps lower than $\bar{E}_1$ are not Pareto-efficient since a marginal increase in $\bar{E}$ results in a Pareto improvement, whereas for caps larger than $\bar{E}_2$, a Pareto improvement can be achieved by reducing $\bar{E}$. Only in the shaded area in the closed interval $[\bar{E}_1,\ \bar{E}_2]$,

[9]Formally, concavity of W_i is shown as part of the proof of Lemma 3 in the Appendix.

we find Pareto-efficient caps. An increase in the emissions cap is detrimental to country 1 here, whereas it benefits country 2.

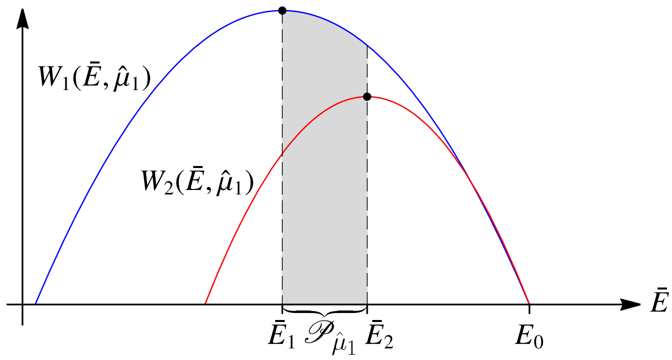

Figure 1: *Individually optimal caps and Pareto set.*

As a next step, we formalize the bargaining procedure. One of the most straightforward and intuitive ways to model a cap negotiation is according to Rubinstein's (1982) alternating-offers model, which applies to our setting as follows.[10]

Country i proposes a cap. Then country $-i$ can either accept this offer and the game ends or reject the offer and make a counteroffer after $\Delta > 0$ time units. In case of rejection, it is i's turn to decide whether to accept the counteroffer or to make a counter-counteroffer. This process continues until one country accepts the proposed cap.[11] A prominent result in bargaining theory is that the subgame perfect equilibrium in the Rubinstein model converges to Nash's (1950) bargaining solution if $\Delta \to 0$ (Binmore, Rubinstein, & Wolinsky, 1986; Binmore, 1987). Intuitively, in the words of Muthoo (1999, p. 52), $\Delta \to 0$ corresponds to a situation where *"the absolute magnitudes of the frictions in the bargaining process are small"*. Evidently, this is in accordance with our setup, as the bargaining process is substantially faster than the underlying process of climate change that requires the reduction of emissions. Even if the bargaining is extended by Δ due to the rejection of an offer, approximately the same benefits and costs can be attained through an agreement in the next round. For simplicity, we assume that the counties have the same time discount rate such that we can apply the symmetric Nash

[10]See Osborne and Rubinstein (1990) and Muthoo (1999) for textbook as well as Roth (1985) and Binmore and Dasgupta (1987) for advanced treatments of bargaining theory.

[11]Note that this standard version of the alternating-offers model does not incorporate the possibility of opting out of the bargaining. In our setup, the interpretation is that, while the counties have already agreed on creating a cap-and-trade system, they only bargain about the implemented cap.

bargaining solution.[12] In our setting, the Nash bargaining solution $\bar{E}^B$ is defined as the solution to the following maximization problem:

$$\max_{\bar{E}} \mathcal{N}_{\hat{\mu}_1}(\bar{E}) := W_1(\bar{E}, \hat{\mu}_1) \cdot W_2(\bar{E}, \hat{\mu}_1), \qquad \text{s.t. } \bar{E} \in \mathscr{P}_{\hat{\mu}_1}, \tag{4.3}$$

where $\mathcal{N}$ is referred to as Nash product.[13] Figure 2 illustrates how the bargaining solution is determined. The purple line represents the Nash product. The Nash bargaining solution is the maximizer of the Nash product among the Pareto-efficient caps, which are represented by the solid part of the purple line.

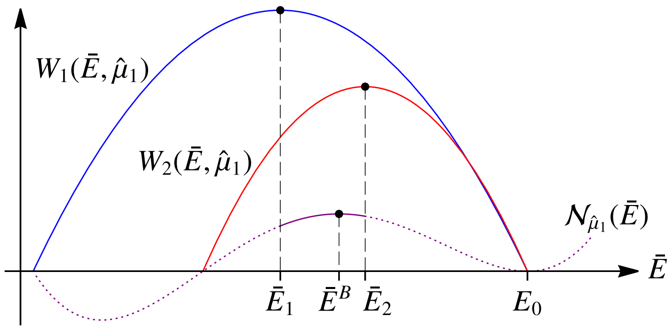

Figure 2: *Nash bargaining solution.*

From differentiation, we obtain the FOC of the Nash product, which reads:

$$\frac{d\mathcal{N}_{\hat{\mu}_1}(\bar{E})}{d\bar{E}} = \sum\nolimits_i \frac{\mathrm{d}W_i(\bar{E}, \hat{\mu}_1)}{\mathrm{d}\bar{E}} \cdot W_{-i}(\bar{E}, \hat{\mu}_1) = 0. \tag{4.4}$$

In our analysis, we exploit the following lemma.

Lemma 4. *There exists a unique Nash bargaining solution* $0 \leq \bar{E}^B < E_0$ *that solves Problem* (4.3). *Any* $\bar{E}^B > 0$ *is determined by the unique solution to* (4.4) *in* $\mathscr{P}_{\hat{\mu}_1}$.

Although (4.4) has multiple solutions (cf. Figure 2), there exists at most one solution that is Pareto-efficient. If that solution indeed exists, then it defines the bargaining outcome $\bar{E}^B > 0$, while otherwise $\bar{E}^B = 0$ holds. Lemma 4 greatly helps us investigate the question of whether bargaining can implement the efficient cap or, more generally, whether the bargaining outcome may be welfare-enhancing

[12] Different discount rates shift bargaining power in favor of country i that possesses a lower discount rate, i.e., that is more patient. This leads to a bargaining outcome that is close to $\bar{E}_i$.

[13] More precisely, $\mathcal{N}_{\hat{\mu}_1}(\bar{E}) = (W_1(\bar{E}, \hat{\mu}_1) - d_1) \cdot (W_2(\bar{E}, \hat{\mu}_1) - d_2)$. As d_i reflects welfare attained by country i if "business as usual" is maintained, it holds that $d_i = 0$ for all $i \in \{1, 2\}$ (Binmore et al., 1986; Muthoo, 1999).

compared to national caps. Since it was shown in Lemma 1 and 3 that $\bar{E}^S > 0$ and $\bar{E}^S \in \mathscr{P}_{\hat{\mu}_1}$, respectively, we can immediately conclude that the bargaining procedure implements the efficient cap if and only if $\bar{E}^S$ solves (4.4).

4.2 Comparison to the Benchmarks

It is worth emphasizing that each total abatement in the joint cap-and-trade system is achieved with the optimal cost structure, namely with equal marginal cost in each country. Hence, if $\bar{E}^B = \bar{E}^S$, then this automatically implies that the bargaining solution leads to the greatest level of overall welfare. Proposition 2 now explores conditions under which bargaining indeed implements $\bar{E}^S$.

Proposition 2 (Bargaining over the Emissions Cap)**.**
Consider two countries that agreed upon setting up a joint cap-and-trade system with an initial allocation $\hat{\mu}_1$ and bargain over the amount of permits. Then the following holds:

(i) *If the countries are sufficiently symmetric in terms of benefits, costs, and initial emissions, then there exists a unique initial allocation $\mu_1^S \in [0, 1]$ for which $\bar{E}^B = \bar{E}^S$.*
(ii) *If, by contrast, $B_i(A(\bar{E}^S)) - B_{-i}(A(\bar{E}^S))$ sufficiently large, then $\bar{E}^B \neq \bar{E}^S$.*

The intuition behind Proposition 2 is due to the efficiency-fairness trade-off in bargaining (see, e.g., Bertsimas, Farias, & Trichakis, 2012; Freeborn, 2023; Dijkstra & Nentjes, 2020). The countries generally face a trade-off between "size of the cake" and "allocation of the cake". On the one hand, they seek to maximize the overall welfare level that they can divide among themselves, i.e., they want to choose a cap close to $\bar{E}^S$. On the other hand, due to equal bargaining power, the countries want to implement a cap that leads to an equal split, i.e., that equalizes individual welfare levels.

Proposition 2 *(i)* reveals that if the countries are sufficiently symmetric in terms of cost- and benefit structures and initial emissions, then a unique allocation $\mu_1^S \in [0, 1]$ exists that completely resolves this trade-off. Given μ_1^S, agreeing on the efficient cap not only maximizes overall welfare but also induces equal welfare levels in both countries. In light of Proposition 1, this implies that if the countries initially agreed on an allocation of permits $\hat{\mu}_1$ sufficiently close to μ_1^S, then determining a joint cap via negotiations is indeed welfare-improving, since

$$W^S \geq W(\bar{E}^B) > W^N,$$

holds by continuity for $\hat{\mu}_1$ sufficiently close to μ_1^S. Put differently, if the countries are sufficiently symmetric and $\hat{\mu}_1$ is in the local neighborhood of μ_1^S, then endogenously designing the cap-and-trade system through cap negotiations alleviates or completely overcomes the free-ride incentive by implementing a stricter emissions cap and enhancing overall welfare compared to national caps. The welfare maximum W^S is however only achieved if the initial allocation $\hat{\mu}_1$ coincides with μ_1^S.

Proposition 2 *(ii)* states that if the countries' benefits obtained under the efficient cap are too different, then bargaining does not implement the social optimum. Intuitively, as indicated by Buchholz, Haupt, and Peters (2005) and Caparrós (2016), the allocation of permits serves as an implicit side payment in an emissions trading system: With a higher share μ_i, country i has to purchase fewer certificates or receives additional revenue for selling the certificates, depending on whether i acts as buyer or seller of permits. However, the scope for providing side payments is limited as country i cannot receive more than the entire share of permits or less than no share. If the countries' benefits obtained under the efficient cap are too different, then the scope for providing implicit side payments is insufficient to fully resolve the trade-off, i.e., there is no allocation of permits for which the efficient cap also yields equal individual welfare levels. Hence, due to the prevailing trade-off, countries forego choosing the efficient cap and instead agree on a cap that leads to a more equal distribution of individual welfare levels.

Since an efficient emissions cap is implemented through an allocation that resolves the trade-off rather than allocating permits proportionally to initial emission levels, we can state the following corollary for grandfathering.

Corollary 1 (Grandfathering)**.**
Consider two countries that agreed upon setting up a joint cap-and-trade system with an initial allocation $\hat{\mu}_1$ and bargain over the amount of permits.
If $\hat{\mu}_1 \propto e_{1,0}/E_0$, then bargaining will generally lead to an inefficient emissions cap.

Example 2 illustrates our results thus far regarding the bargaining outcome.

Example 2 (Optimal Initial Allocation).
Revisit the symmetric parameterization of Example 1. Figure 3 depicts the initial allocation μ_1^S *for different* β_2 *and* $\beta_1 = 0.3$*. For completely symmetric countries,* $(\beta_2/\beta_1 = 1)$*, bargaining implements the social optimum if and only if the countries initially agreed on an equal distribution of permits,* $\hat{\mu}_1 = 0.5$*. By contrast, for* $\beta_2/\beta_1 < 0.8$ *and* $\beta_2/\beta_1 > 1.2$*, the means of implicit side payments are insufficient to implement the efficient cap. Grandfathering does not implement the efficient cap except for the special case where* $\beta_2/\beta_1 = 1$*.*

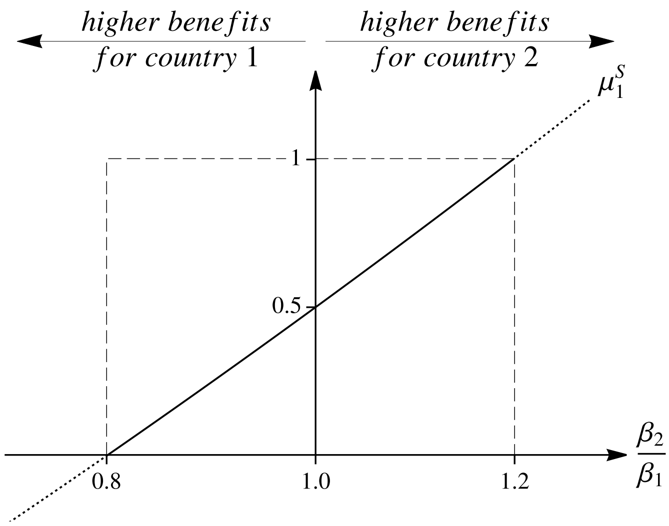

Figure 3: *Optimal initial allocation.*
($e_{1,0} = e_{2,0} = 0.5$*,* $\zeta_1 = \zeta_2 = 0.5$*, and* $\beta_1 = 0.3$*)*

5 Negotiations on Cap & Allocation

Equipped with the inside that bargaining can be welfare-improving, we now add a further degree of freedom by allowing countries to negotiate simultaneously on the emissions cap and the initial allocation of permits. We maintain the assumption that countries cannot strategically end the negotiation for the time being but will abandon this assumption in the course of this section. The Pareto set $\mathscr{P}$ now consists of tuples $(\bar{E}, \mu_1)$, i.e., combinations of a joint emissions cap and a corresponding allocation of these certificates among the countries. It can be defined as follows.

Definition 2 (Pareto-Efficient Tuples)**.**
The set of Pareto-efficient tuples, $\mathscr{P} \subset \mathbb{R}_+ \times [0,1]$, *is the set of all tuples* $(\bar{E}, \mu_1)$ *for which no other tuple* $(\bar{E}', \mu_1')$ *exists that satisfies*

$$W_i(\bar{E}', \mu_1') \geq W_i(\bar{E}, \mu_1), \quad \textit{and} \quad W_{-i}(\bar{E}', \mu_1') > W_{-i}(\bar{E}, \mu_1)$$

for at least one $i \in \{1, 2\}$, *where* $\mu_1, \mu_1' \in [0, 1]$.

For the bargaining procedure, this implies that an offer in the alternating-offers model now consists of a tuple $(\bar{E}, \mu_1)$, i.e., a proposal about the amount of permits and their allocation among the countries. Due to the negligible friction in the bargaining process, we can exploit the relation between the subgame perfect equilibrium in this dynamic game and the static Nash bargaining approach again. Indeed, the bargaining solution $(\bar{E}^B, \mu_1^B)$ is defined by a solution to the following maximization problem:

$$\max_{(\bar{E}, \mu_1)} \mathcal{N}(\bar{E}, \mu_1) \coloneqq W_1(\bar{E}, \mu_1) \cdot W_2(\bar{E}, \mu_1) \qquad \text{s.t. } (\bar{E}, \mu_1) \in \mathscr{P}. \tag{5.1}$$

Analyzing the Problem (5.1) leads to the following proposition.

Proposition 3 (Bargaining over the Emissions Cap and Initial Allocation)**.**
Consider two countries that agreed upon setting up a joint cap-and-trade system and bargain over the amount and allocation of permits. Then the following holds:

(i) If the countries are sufficiently symmetric in terms of benefits, costs, and initial emissions, then $\bar{E}^B = \bar{E}^S$ *and* $\mu_1^B = \mu_1^S$.
(ii) If, by contrast, $B_i(A(\bar{E}^S)) - B_{-i}(A(\bar{E}^S))$ *sufficiently large, then* $\bar{E}^B \neq \bar{E}^S$.

Allowing the countries to negotiate simultaneously on the allocation of permits and the emissions cap yields some interesting results. Proposition 3 *(i)* shows that whenever the scope for setting side payments allows the countries to resolve the trade-off and implement the efficient cap, they will, in fact, design an allocation of permits to do so.[14] Put differently, if the countries are sufficiently symmetric, then setting up a cap-and-trade system and letting the countries bargain over the amount and allocation of permits completely removes the distortions created by the free-riding incentive. It is worth emphasizing that the countries themselves then design an emissions cap exactly as it would have been done by a social planner or regulator with complete information.

[14] Technically, the following relation holds $\bar{E}^B = \bar{E}^S \iff \mu_1^B = \mu_1^S \iff \mu_1^S \in [0, 1]$.

Proposition 3 *(ii)* is due to the mechanism that we have already encountered: If the countries' benefits from the efficient cap are too heterogeneous, then the scope for providing side payments is insufficient to implement the efficient cap. This holds true irrespective of whether the allocation of permits is exogenously given or endogenously determined via bargaining.

It is worth discussing our results against the background of the famous COASE THEOREM. As summarized by Harris and Roach (2022, p. 60), the theorem states that *"if property rights are well defined, and no significant transaction costs exist, an efficient allocation of resources will result even with externalities."* In our model, an efficient cap obtains, except for the case where the scope for side payments is exhausted, which can be interpreted as infinite transaction costs. While our results are thus in line with the COASE THEOREM, it should be stressed that they are *not* a mere consequence of the theorem. First, property rights in our setup are ex ante not well defined, as the exact purpose of the negotiation is to determine these rights by specifying the amount of permits and their allocation among the countries. Second, even if this prerequisite was satisfied, Hahnel and Sheeran (2009) argue that formal bargaining models will not generally result in the efficient outcome predicted by the rather informal COASE THEOREM: Whether negotiations lead to an efficient outcome crucially depends on the bargaining procedure, countries' welfare functions and time preferences, as well as the information structure in the game.

The preceding analysis focused exclusively on situations where the countries were unable to terminate the negotiation strategically. While plausible for some settings, others may allow each side to strategically opt out and end the bargaining in disagreement. Metaphorically speaking, if one party decides to leave the negotiation table, then both parties are left with their *outside option*. Since the countries must rely on national caps in case setting up a joint cap-and-trade system fails, they are left with the Nash equilibrium caps described in Lemma 2. Hence, country i's outside option is simply W_i^N, i.e., the welfare level obtained in the national cap benchmark. As the friction in the bargaining process is negligibly small, insights from bargaining theory allow us to link the Nash bargaining solution to subgame perfect equilibrium in the alternating offers model extended by the possibility of strategic opting out for both parties. Following Binmore (1985); Binmore et al. (1986), and Muthoo (1999), the subgame perfect equilibrium in the extended alternating-offers model converges to the solution to the following maximization problem:

$$\begin{aligned} &\max_{\bar{E},\mu_1} \mathcal{N}(\bar{E},\mu_1) = W_1(\bar{E},\mu_1) \cdot W_2(\bar{E},\mu_1) \\ \text{s.t.} \quad &(\bar{E},\mu_1) \in \mathscr{P}, \quad W_1(\bar{E},\mu_1) \geq W_1^N, \quad W_2(\bar{E},\mu_1) \geq W_2^N. \end{aligned} \tag{5.2}$$

The only difference to Problem (5.1) is that each country can secure itself a welfare level weakly greater than its outside option. Intuitively, a country would never accept a "bad" offer in the negotiation but instead strategically opt out and realize its outside option. Examining Problem (5.2) yields the following proposition.

Proposition 4 (Bargaining with Outside Option)**.**
Consider two countries that bargain on setting up a joint cap-and-trade system. If the countries are sufficiently symmetric in terms of benefits, costs and initial emissions, then they agree on setting up a joint cap-and-trade system with $\bar{E}^B = \bar{E}^S$ *and* $\mu_1^B = \mu_1^S$*.*

Proposition 4 states if two countries are sufficiently symmetric, instead of implementing national caps, they will agree on setting up a joint cap-and-trade system with an emissions cap that is efficient. This leads to the highest possible level of welfare, which is equally distributed among the countries via the allocation of permits. Our analysis, therefore, points out a simple way to circumvent the inefficiency caused by the free-riding incentive: Countries should be enabled to set up joint cap-and-trade systems and allowed to negotiate the cap and allocation of permits. This procedure then implements the social planner result, provided that the countries are sufficiently symmetric. Otherwise, the free-riding incentive might be so strong for a country and its outside option thus so attractive that the negotiated joint emissions cap is distorted away from the social optimum. This is illustrated in the following example.

Example 3 (Bargaining with Outside Option).
Revisit the symmetric parameterization of Example 1, where $\beta_1 = 0.3$.

(i) *For* $\beta_2/\beta_1 \in [0.8, 0.97)$ *bargaining implements* $\bar{E}^S$*, while* μ_1 *is determined by providing a welfare level to country 1 that equals its outside option.*[15]

(ii) *For* $\beta_2/\beta_1 \in [0.97, 1.03]$ *bargaining implements* $\bar{E}^S$ *and* μ_1 *is determined by equalizing the corresponding welfare levels in both countries.*

(iii) *For* $\beta_2/\beta_1 \in (1.03, 2.05]$ *bargaining implements* $\bar{E}^S$*, while* μ_1 *is determined by providing a welfare level to country 2 that equals its outside option.*

(iv) *By contrast, for* $\beta_2/\beta_1 > 2.05$ *bargaining does not implement* $\bar{E}^S$*.*

The example shows that the problem of insufficient side payments carries over to the presence of outside options. Even more surprisingly, comparing Examples 2 and 3 reveals that the presence of an outside option may ensure that the bargaining leads to an efficient cap for more asymmetric countries than it would be the case without an outside option. This is precisely the case for β_2/β_1 ratios between 1.2 and 2.05. We summarize this surprising finding in the following proposition.

Proposition 5 (Presence vs. Absence of an Outside Option).
If countries have the possibility of strategic opting out, then bargaining may implement an efficient cap for more asymmetric countries than it would be the case without an outside option.

The intuition is as follows. If country i's outside option is sufficiently attractive, then this eliminates the efficiency-fairness trade-off since the welfare level granted to country i equals its outside option. Whenever possible, the parties then agree on the efficient cap and an allocation of permits that provides county i with the welfare level of its outside option.[16] This bargaining outcome not only ensures county i's participation in the scheme but also maximizes the welfare left for country $-i$.

[15] We do not examine ratios $\beta_2/\beta_1 < 0.8$, as they violate Assumption 2.

[16] However, as can be seen in Part *(iv)* of Example 3, if the countries are too different, then there may not be a feasible allocation of the efficient amount of permits that provides country i with the welfare level of its outside option. In this case, the countries will not agree on the efficient cap but rather on one closer to country i's individually optimal cap $\bar{E}_i$.

6 Multilateral Negotiations

Although our analysis was conducted in the two-country case for the sake of clarity, it easily generalizes to the case of n countries, where $n > 2$. Analogously to the two-country case without an outside option, bargaining implements the efficient cap if and only if a *feasible* allocation of permits exists that equates the welfare levels in all countries resulting from the efficient cap.[17] A feasible allocation is now characterized by

$$\sum_i \mu_i = 1, \qquad \text{and} \qquad \mu_i \in [0,1], \qquad\qquad i = 1, \ldots, n.$$

In the symmetric case, allocating equal shares $\mu_i = 1/n$ to all countries n indeed constitutes the unique feasible allocation that equates the welfare levels from the efficient cap such that bargaining leads to a design of the cap-and-trade system that reduces emissions to the efficient level. Hence, continuity implies that bargaining also implements the efficient cap if the countries are sufficiently symmetric. If, by contrast, $B_i(A(\bar{E}^S)) - B_j(A(\bar{E}^S))$ is sufficiently large for at least two countries i and j, then equating welfare levels resulting from the efficient cap would either require $\mu_i < 0$ or $\mu_j > 1$ such that no feasible allocation exists to do so. Bargaining will therefore not implement the efficient cap in this case. In the presence of an outside option, it additionally holds for symmetric countries that

$$W_i(\bar{E}^S, 1/n) = \frac{W^S}{n} > \frac{W^N}{n} = W_i^N, \qquad\qquad i = 1, \ldots, n.$$

We can thus infer that even with an outside option if the n countries are sufficiently symmetric, they will agree on setting up a cap-and-trade system and cap the emissions at the efficient level. Accordingly, our analysis carries over entirely to the case with more than two countries, i.e., multilateral negotiations on the design of a joint cap-and-trade system.

7 Conclusion

How can global greenhouse gas emissions be reduced to mitigate climate change? We have addressed this question by analyzing whether designing an emissions trading scheme through negotiation has the potential to enforce efficient emissions

[17] Technically, the FOC of the social planner and the FOC of the Nash product with respect to the emissions cap coincide in this case for the efficient cap.

levels. Our analysis builds on a simple model with two countries that experience a positive externality from reducing emissions. Due to this externality, each country has an incentive to free-ride on the other country's abatement activities. In the case of national abatement activities, free-riding leads to an overall emissions level that exceeds the social optimum.

By applying insights from bargaining theory, we find that the ecological market failure resulting from the free-riding incentive may be eliminated by a simple mechanism derived from quantity-based instruments: Enabling the countries to set up a joint cap-and-trade system and allowing them to bargain over the amount and allocation of certificates. If the countries are sufficiently symmetric, they agree to cap emissions at the efficient level. Since this efficient emissions level obtains with the optimal distribution of abatement activities among the countries, i.e., at the lowest cost, the endogenous cap maximizes overall welfare. The countries then use the allocation of certificates as an implicit side payment to distribute welfare equally among themselves. Surprisingly, an efficient cap may also be achieved for even more asymmetric countries if they can strategically terminate the bargaining and deploy national caps instead. However, if the countries are too different, then bargaining may not necessarily result in the efficient cap. In this case, the scope for implicit side payments through the initial allocation of certificates may be insufficient to make both countries agree on the efficient cap.

The implications of our analysis are quite striking. Even in the absence of a social planner, a joint cap-and-trade system may induce efficient outcomes. The sheer possibility of negotiating its design then induces cooperative behavior: The countries overcome the free-riding incentives, implement the efficient emissions cap, and distribute the resulting overall welfare among themselves in a fair way. Our results imply that negotiations are pivotal in efficiently designing cap-and-trade systems and should thus be encouraged. Moreover, they underline the importance of removing all sorts of barriers, such as transaction costs (cf. Montero, 1998), imperfections in the emissions market (cf. Stavins, 2008b), and conflicting national regulations (cf. Hahn & Stavins, 2011), that either hinder countries from setting up a joint cap-and-trade system or prevent them from linking existing schemes. However, designing a mechanism that implements efficient outcomes for strongly asymmetric countries, especially in the presence of outside options, is significantly more complex and constitutes a fruitful avenue for further research.

References

Aldy, J. E., Barrett, S., & Stavins, R. N. (2003). Thirteen plus one: A comparison of global climate policy architectures. *Climate Policy*, *3*(4), 373–397.

Arvaniti, M., & Habla, W. (2021). The political economy of negotiating international carbon markets. *Journal of Environmental Economics and Management*, *110*, 102521.

Barrett, S. (1994). Self-enforcing international environmental agreements. *Oxford Economic Papers*, *46*(Supplement 1), 878–894.

Barrett, S. (2003). *Environment and statecraft: The strategy of environmental treaty-making.* Oxford: Oxford University Press.

Baudry, M., Faure, A., & Quemin, S. (2021). Emissions trading with transaction costs. *Journal of Environmental Economics and Management*, *108*, 102468.

Bertsimas, D., Farias, V. F., & Trichakis, N. (2012). On the efficiency-fairness trade-off. *Management Science*, *58*(12), 2234–2250.

Binmore, K. (1985). Bargaining and coalitions. In A. Roth (Ed.), *Game-theoretic models of bargaining* (p. 269-–304). Cambridge, UK: Cambridge University Press.

Binmore, K. (1987). Nash bargaining theory ii. In K. Binmore & P. Dasgupta (Eds.), *The economics of bargaining* (pp. 61–76). Oxford: Blackwell.

Binmore, K., & Dasgupta, P. (1987). *The economics of bargaining.* Oxford: Blackwell.

Binmore, K., Rubinstein, A., & Wolinsky, A. (1986). The nash bargaining solution in economic modelling. *RAND Journal of Economics*, *17*(2), 176–188.

Böhringer, C., & Lange, A. (2005). On the design of optimal grandfathering schemes for emission allowances. *European Economic Review*, *49*(8), 2041–2055.

Buchholz, W., Haupt, A., & Peters, W. (2005). International environmental agreements and strategic voting. *Scandinavian Journal of Economics*, *107*(1), 175–195.

Caparrós, A. (2016). Bargaining and international environmental agreements. *Environmental and Resource Economics*, *65*(1), 5–31.

Caparrós, A., Péreau, J.-C., & Tazdaït, T. (2004). North-south climate change negotiations: A sequential game with asymmetric information. *Public Choice*, *121*(3–4), 455–480.

Carraro, C., & Siniscalco, D. (1993). Strategies for the international protection of the environment. *Journal of Public Economics*, *52*(3), 309–328.

Coase, R. H. (1960). The problem of social cost. *Journal of Law and Economics*, *3*, 1–44.

Dales, J. H. (1968). *Pollution, property & prices; an essay in policy-making and economics.* Toronto: University of Toronto Press.

Damon, M., Cole, D. H., Ostrom, E., & Sterner, T. (2019). Grandfathering: Environmental uses and impacts. *Review of Environmental Economics and Policy*, *13*(1), 23–42.

Dijkstra, B. R., & Nentjes, A. (2020). Pareto-efficient solutions for shared public good provision: Nash bargaining versus exchange-matching-lindahl. *Resource and Energy Economics*, *61*, 101179.

Doda, B., Quemin, S., & Taschini, L. (2019). Linking permit markets multilaterally. *Journal of Environmental Economics and Management*, *98*, 102259.

Doda, B., & Taschini, L. (2017). Carbon dating: When is it beneficial to link etss? *Journal of the Association of Environmental and Resource Economists*, *4*(3), 701–730.

Egenhofer, C. (2007). The making of the eu emissions trading scheme: Status, prospects and implications for business. *European Management Journal*, *25*(6), 453–463.

Epstein, R. A. (1979). Possession as the root of title. *Georgia Law Review*, *13*, 1221—43.

Flachsland, C., Marschinski, R., & Edenhofer, O. (2009). To link or not to link: benefits and disadvantages of linking cap-and-trade systems. *Climate Policy*, *9*(4), 358–372.

Freeborn, D. (2023). Efficiency and fairness trade-offs in two player gargaining games. *European Journal for Philosophy of Science*, *13*(49), 1–23.

Gersbach, H., & Hummel, N. (2016). A development-compatible refunding scheme for a climate treaty. *Resource and Energy Economics*, *44*, 139–168.

Goulder, L. H., & Schein, A. R. (2013). Carbon taxes versus cap and trade: A critical review. *Climate Change Economics*, *4*(3), 1350010.

Graziosi, G. R. (2009). On the strategic use of representative democracy in international agreements. *Journal of Public Economic Theory*, *11*(2), 281–296.

Grimm, V., & Ilieva, L. (2013). An experiment on emissions trading: The effect of different allocation mechanisms. *Journal of Regulatory Economics*, *44*(3), 308–338.

Habla, W., & Winkler, R. (2018). Strategic delegation and international permit markets: Why linking may fail. *Journal of Environmental Economics and Management*, *92*, 244–250.

Hahn, R. W. (1984). Market power and transferable property rights. *Quarterly Journal of Economics*, *99*(4), 753–765.

Hahn, R. W., & Stavins, R. N. (2011). The effect of allowance allocations on cap-and-trade system performance. *Journal of Law and Economics*, *54*(4), S267–S294.

Hahnel, R., & Sheeran, K. A. (2009). Misinterpreting the coase theorem. *Journal of Economic Issues*, *43*(1), 215–238.

Hardin, G. (1968). The tragedy of the commons. *Science*, *162*(3859), 1243–1248.

Harris, J. M., & Roach, B. (2022). *Environmental and natural resource economics: A contemporary approach.* New York: Routledge.

Harstad, B. (2007). Harmonization and side payments in political cooperation. *American Economic Review*, *97*(3), 871–889.

Harstad, B. (2023). Pledge-and-review bargaining: From kyoto to paris. *Economic Journal*, *133*(651), 1181–1216.

Helm, C. (2003). International emissions trading with endogenous allowance choices. *Journal of Public Economics*, *87*(12), 2737–2747.

IPCC. (2023). Summary for policymakers. *Climate Change 2023: Synthesis Report. Contribution of Working Groups I, II and III to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*, 1–34.

Lueck, D. (1995). The rule of first possession and the design of the law. *Journal of Law and Economics*, *38*(2), 393–436.

Mas-Colell, A., Whinston, M. D., & Green, J. R. (1995). *Microeconomic theory.* Oxford: Oxford University Press.

McGinty, M. (2007). International environmental agreements among asymmetric nations. *Oxford Economic Papers*, *59*(1), 45–62.

Montero, J.-P. (1998). Marketable pollution permits with uncertainty and transaction costs. *Resource and Energy Economics*, *20*(1), 27–50.

Muthoo, A. (1999). *Bargaining theory with applications.* Cambridge, UK: Cambridge University Press.

Nash, J. F. (1950). The bargaining problem. *Econometrica*, *18*(2), 155–162.

Nehra, A., & Caplan, A. J. (2022). Nash bargaining in a general equilibrium framework: The case of a shared surface water supply. *Water Resources and Economics*, *39*, 100206.

Nordhaus, W. D. (2007). To tax or not to tax: Alternative approaches to slowing global warming. *Review of Environmental Economics and Policy*, *1*(1), 26–44.

Nordhaus, W. D. (2015). Climate clubs: Overcoming free-riding in international climate policy. *American Economic Review*, *105*(4), 1339–1370.

Nordhaus, W. D. (2019). Climate change: The ultimate challenge for economics. *American Economic Review*, *109*(6), 1991–2014.

Osborne, M. J., & Rubinstein, A. (1990). *Bargaining and markets.* San Diego: Academic Press.

Pigou, A. (1920). *The economics of welfare.* London: McMillan & Co.

Rose, C. M. (1985). Possession as the origin of property. *University of Chicago Law Review*, *52*(1), 73–88.

Roth, A. E. (1985). *Game-theoretic models of bargaining.* Cambridge, UK: Cambridge University Press.

Rubinstein, A. (1982). Perfect equilibrium in a bargaining model. *Econometrica*, *50*(1), 97–109.

Schmalensee, R., & Stavins, R. N. (2017). Lessons learned from three decades of experience with cap and trade. *Review of Environmental Economics and Policy*, *11*(1), 59–79.

Segendorff, B. (1998). Delegation and threat in bargaining. *Games and Economic Behavior*, *23*(2), 266–283.

Smead, R., Sandler, R. L., Forbes, P., & Basl, J. (2014). A bargaining game analysis of international climate negotiations. *Nature Climate Change*, *4*(6), 442–445.

Stavins, R. N. (1995). Transaction costs and tradeable permits. *Journal of Environmental Economics and Management*, *29*(2), 133–148.

Stavins, R. N. (2008a). Addressing climate change with a comprehensive us cap-and-trade system. *Oxford Review of Economic Policy*, *24*(2), 298–321.

Stavins, R. N. (2008b). A meaningful u.s. cap-and-trade system to address climate change. *Harvard Environmental Law Review*, *32*(2), 293–371.

Stavins, R. N. (2022). The relative merits of carbon pricing instruments: Taxes versus trading. *Review of Environmental Economics and Policy*, *16*(1), 62–82.

Stern, N. H. (2007). *The economics of climate change: The stern review.* Cambridge, UK: Cambridge University Press.

Stern, N. H. (2008). The economics of climate change. *American Economic Review: Papers & Proceedings*, *98*(2), 1–37.

Swanson, T., & Groom, B. (2012). Regulating global biodiversity: What is the oroblem? *Oxford Review of Economic Policy*, *28*(1), 114–138.

Underdal, A., Hovi, J., Kallbekken, S., & Skodvin, T. (2012). Can conditional commitments break the climate change negotiations deadlock? *International Political Science Review*, *33*(4), 475–493.

Urpelainen, J. (2012). Technology investment, bargaining, and international environmental agreements. *International Environmental Agreements: Politics, Law and Economics*, *12*, 145–163.

Weitzman, M. L. (1974). Prices vs. quantities. *Review of Economic Studies*, *41*(4), 477–491.

Weitzman, M. L. (2007). A review of the stern review on the economics of climate change. *Journal of Economic Literature*, *45*(3), 703–724.

Weitzman, M. L. (2014). Can negotiating a uniform carbon price help to internalize the global warming externality? *Journal of the Association of Environmental and Resource Economists*, *1*(1/2), 29–49.

Woerdman, E., Arcuri, A., & Clò, S. (2008). Emissions trading and the polluter-pays principle: Do polluters pay under grandfathering? *Review of Law & Economics*, *4*(2), 565–590.

A Appendix: Proofs

Proof of Lemma 1. First, we establish the existence of a solution to (3.3). Since

$$\lim_{\bar{E}\to 0} \sum_i B_i'(A(\bar{E})) - C_i'(a_i(\bar{E})) = -\zeta_1\zeta_2 E_0 < 0,$$

$$\lim_{\bar{E}\to E_0} \sum_i B_i'(A(\bar{E})) - C_i'(a_i(\bar{E})) = \sum_i B_i'(0) = \infty > 0,$$

the Intermediate Value Theorem immediately implies the existence of a solution. Moreover, differentiation with respect to $\bar{E}$ yields

$$\sum_i B_i''(A(\bar{E}))\frac{\partial A(\bar{E})}{\partial \bar{E}} - C_i''(a_i(\bar{E}))\frac{\partial a_i(\bar{E})}{\partial \bar{E}},$$

which using (2.7)–(2.8), simplifies to

$$-\sum_i B_i''(A(\bar{E})) + \zeta_1\zeta_2 > 0.$$

Hence, the solution is unique. Note also that our assumptions regarding the functional form of the cost and benefit functions immediately imply that $W(\bar{E})$ is strictly concave. Therefore, the solution $\bar{E}^S$ to the FOC is indeed a maximizer.

□

Proof of Lemma 2. In the Nash equilibrium, both countries choose national caps that are best responses to each other. To derive the best responses, note that country i's objective function in (3.6) is concave in $\bar{e}_i$ such that if a solution to the FOC exists, then the solution indeed maximizes country i's welfare.

We consider the best response of country 1 first by investigating the limits of its FOC for a given $\bar{e}_2$

$$\lim_{\bar{e}_1\to e_{1,0}} B_1'(A(\bar{e}_1,\bar{e}_2)) - C_1'(a_1(\bar{e}_1)) = B_1'(A(e_{1,0},\bar{e}_2)) > 0,$$

$$\lim_{\bar{e}_1\to 0} B_1'(A(\bar{e}_1,\bar{e}_2)) - C_1'(a_1(\bar{e}_1)) = B_1'(A(0,\bar{e}_2)) - \zeta_1 e_{1,0} \gtreqless 0.$$

Now, we need to distinguish two cases:

Case 1. If $B_1'(A(0,\bar{e}_2)) - \zeta_1 e_{1,0} \geq 0$, then $\bar{e}_1^N = 0$. Since, for country 2, it holds that

$$\lim_{\bar{e}_2\to e_{2,0}} B_2'(A(\bar{e}_2,0)) - C_2'(a_2(\bar{e}_2)) = B_2'(A(e_{2,0},0)) > 0,$$

$$\lim_{\bar{e}_2\to 0} B_2'(A(\bar{e}_2,0)) - C_2'(a_2(\bar{e}_2)) = -\zeta_2 e_{2,0} < 0.$$

The INTERMEDIATE VALUE THEOREM implies the existence of a solution $0 < \bar{e}_2^N < e_{2,0}$ to country 2's FOC. Moreover, differentiation with respect to $\bar{e}_2$ and simplifying yields

$$-B_2''(A(\bar{e}_2, 0)) + C_2''(a_2(\bar{e}_2)) = -B_2''(A(\bar{e}_2, 0)) + \zeta_2 > 0.$$

Hence, $\bar{e}_2^N$ is unique. In this case, the Nash equilibrium caps are thus uniquely determined and satisfy $\bar{e}_1^N = 0$ and $0 < \bar{e}_2^N < e_{2,0}$.

Case 2. If $B_1'(A(0, \bar{e}_2)) - \zeta_1 e_{1,0} > 0$, then the INTERMEDIATE VALUE THEOREM implies the existence of a solution to country 1's FOC. Again, differentiating with respect to $\bar{e}_1$ and simplifying yields

$$-B_1''(A(\bar{e}_1, \bar{e}_2)) + C_1''(a_1(\bar{e}_1)) = -B_1''(A(\bar{e}_1, \bar{e}_2)) + \zeta_1 > 0.$$

such that the FOC has a unique solution that defines a best response function of the form $\bar{e}_1(\bar{e}_2)$. Plugging $\bar{e}_1(\bar{e}_2)$ into country 2's FOC and analyzing the limits yields

$$\lim_{\bar{e}_2 \to e_{2,0}} B_2'(A(\bar{e}_2, \bar{e}_1(\bar{e}_2))) - C_2'(a_2(\bar{e}_2)) = B_2'(A(e_{2,0}, \bar{e}_1(e_{2,0}))) > 0,$$

$$\lim_{\bar{e}_2 \to 0} B_2'(A(\bar{e}_2, \bar{e}_1(\bar{e}_2))) - C_2'(a_2(\bar{e}_2)) = B_2'(A(0, \bar{e}_1(0))) - \zeta_2 e_{2,0} \gtreqless 0.$$

Two subcases need to be distinguished now:

Case 2.1. If $B_2'(A(0, \bar{e}_1(0))) - \zeta_2 e_{2,0} \geq 0$, then $\bar{e}_2^N = 0$. Again, the Nash equilibrium caps are unique and satisfy $0 < \bar{e}_1^N = \bar{e}_1(0) < e_{1,0}$ and $\bar{e}_2^N = 0$.

Case 2.2. If $B_2'(A(0, \bar{e}_1(0))) - \zeta_2 e_{2,0} < 0$, then, by the INTERMEDIATE VALUE THEOREM, there exists a solution to country 2's FOC. Differentiation with respect to $\bar{e}_2$ and simplifying yields

$$-B_2''(A(\bar{e}_2, \bar{e}_1(\bar{e}_2))) \left[1 + \frac{\partial \bar{e}_1(\bar{e}_2)}{\partial \bar{e}_2}\right] + \zeta_2 \tag{A.1}$$

From differentiating country 1's FOC and simplifying, we get that

$$\frac{\partial \bar{e}_1(\bar{e}_2)}{\partial \bar{e}_2} = \frac{B_1''(A(\bar{e}_1, \bar{e}_2))}{-B_1''(A(\bar{e}_1, \bar{e}_2)) + \zeta_1} \in (-1, 0) \qquad \text{for } 0 < \bar{e}_1 < e_{1,0}. \tag{A.2}$$

In view of (A.2), we find that the term in (A.1) is strictly positive. Hence, the solution to country 2's FOC is unique. Therefore, the unique Nash equilibrium caps in this case are $0 < \bar{e}_1^N = \bar{e}_1(\bar{e}_2^N) < e_{1,0}$ and $0 < \bar{e}_2^N < e_{2,0}$. □

Proof of Proposition 1. We distinguish two cases depending on whether $0 < \bar{e}_1^N, \bar{e}_2^N$ or $0 < \bar{e}_i^N$ and $0 = \bar{e}_{-i}^N$. For both cases, it is shown that $\bar{E}^S < \sum_i \bar{e}_i^N$ holds.

Case 1. We start by considering the case where $0 < \bar{e}_1^N, \bar{e}_2^N$. In this case, both FOCs in (3.7) hold with equality for the Nash equilibrium caps. Summing up these equations, we obtain from (3.7) that

$$\sum_i B_i'(A(\bar{e}_i^N, \bar{e}_{-i}^N)) = \sum_i C_i'(a_i(\bar{e}_i^N)) \tag{A.3}$$

must hold. Assume now, for the sake of contradiction, that $\bar{E}^S \geq \sum_i \bar{e}_i^N$ would hold. This implies that

$$\sum_i B_i'(A(\bar{e}_i^N, \bar{e}_{-i}^N)) \leq \sum_i B_i'(A(\bar{E}^S)) = C_i'(a_i(\bar{E}^S)) < \sum_i C_i'(a_i(\bar{e}_i^N)), \tag{A.4}$$

where the last inequality follows from the fact that $a_i(\bar{e}_i^N) \geq a_i(\bar{E}^S)$ as well as $a_{-i}(\bar{e}_{-i}^N) > 0$ must hold for at least one $i \in \{1, 2\}$ if $\bar{E}^S \geq \sum_i \bar{e}_i^N$. Comparing (A.3) and (A.4) yields a contradiction. We must thus have that $\bar{E}^S < \sum_i \bar{e}_i^N$.

Case 2. Now, consider the case where $0 < \bar{e}_i^N < e_{i,0}$ and $0 = \bar{e}_{-i}^N$. In this case, country i's FOCs in (3.7) holds with equality for the Nash equilibrium caps

$$B_i'(A(\bar{e}_i^N, 0)) = C_i'(a_1(\bar{e}_i^N)). \tag{A.5}$$

Assume, for the sake of contradiction, that $\bar{E}^S \geq \sum_i \bar{e}_i^N = \bar{e}_i^N$ would hold. This implies that

$$B_i'(A(\bar{e}_i^N, 0)) \leq B_i'(A(\bar{E}^S)) < \sum_i B_i'(A(\bar{E}^S)) = C_i'(a_i(\bar{E}^S)) \leq C_i'(a_i(\bar{e}_i^N)), \tag{A.6}$$

where the last inequality follows from the fact that $a_i(\bar{e}_i^N) \leq a_i(\bar{E}^S)$ must hold if $\bar{E}^S \geq \bar{e}_i^N$. Comparing (A.5) and (A.6) yields a contradiction. We must thus have that $\bar{E}^S < \sum_i \bar{e}_i^N$.

To see that $\bar{E}^S$ indeed induces a higher level of overall welfare, simply note that

$$W^S = W(\bar{E}^S) > W(\textstyle\sum_i \bar{e}_i^N) \geq W^N. \tag{A.7}$$

The first inequality follows from the fact that $\bar{E}^S < \sum_i \bar{e}_i^N$ where $\bar{E}^S$ is the unique maximizer of W. The second inequality holds since the cap and trade system realizes the benefits from capping the overall emission to the level $\sum_i \bar{e}_i^N$ at the lowest possible cost, i.e., a distribution of abatement activities among the countries that equates their marginal costs. □

Proof of Lemma 3. First, we show that Problem (4.1) is strictly concave. From differentiating the l.h.s. of (4.2), we get

$$\frac{\mathrm{d}^2 W_i(\bar{E},\hat{\mu}_1)}{\mathrm{d}\bar{E}^2} = -\,B_i''(A(\bar{E}))\frac{\partial A(\bar{E})}{\partial \bar{E}} + C_i''(a_i(\bar{E}))\frac{\partial a_i(\bar{E})}{\partial \bar{E}}\left(2\mu_i + \frac{\partial a_i(\bar{E})}{\partial \bar{E}}\right)$$

$$\text{where} \quad \mu_1 = \hat{\mu}_1, \quad \text{and} \quad \mu_2 = 1-\hat{\mu}_1,$$

since $C_i''' = \partial^2 a_i(\bar{E})/\partial \bar{E}^2 = 0$. Using (2.7), (2.8), and the fact that any $\mu_i \in [0,1]$, we obtain

$$\begin{aligned} \frac{\mathrm{d}^2 W_i(\bar{E},\hat{\mu}_1)}{\mathrm{d}\bar{E}^2} &= B_i''(A(\bar{E})) + C_i''(a_i(\bar{E}))\Big(-2\mu_i\zeta_{-i} + \zeta_{-i}^2\Big) \\ &\leq B_i''(A(\bar{E})) + C_i''(a_i(\bar{E}))\zeta_{-i}^2 \\ &= B_i''(A(\bar{E})) + \zeta_i\zeta_{-i} < 0. \end{aligned} \tag{A.8}$$

Next, we investigate whether the FOC has a solution by considering the limits

$$\begin{aligned} &\lim_{\bar{E}\to E_0} -B_i'(A(\bar{E})) + \mu_i C_i'(a_i(\bar{E})) + x_{i,-i}(\bar{E})\; C_i''(a_i(\bar{E}))\frac{\partial a_i(\bar{E})}{\partial \bar{E}} = -\infty < 0, \\ &\lim_{\bar{E}\to 0} -B_i'(A(\bar{E})) + \mu_i C_i'(a_i(\bar{E})) + x_{i,-i}(\bar{E})\; C_i''(a_i(\bar{E}))\frac{\partial a_i(\bar{E})}{\partial \bar{E}} \\ &\qquad\qquad = \zeta_1\zeta_2 E_0\Big(\mu_i - \zeta_{-i} + \frac{e_{i,0}}{E_0}\Big) \gtreqless 0. \end{aligned}$$

Now, we need to distinguish two cases:

Case 1. If $\mu_i - \zeta_{-i} + \frac{e_{i,0}}{E_0} > 0$, then the Intermediate Value Theorem implies the existence of solution $0 < \bar{E}_i < E_0$ to the FOC (4.2). Since Problem (4.1) is strictly concave, this solution must be unique.

Case 2. If $\mu_i - \zeta_{-i} + \frac{e_{i,0}}{E_0} \leq 0$, then strict concavity of Problem (4.1) immediately implies that $\bar{E}_i = 0$. Using that $\zeta_{-i} = 1-\zeta_i$, $e_{i,0} = E_0 - e_{-i,0}$ and $\mu_{-i} = 1-\mu_i$, it holds for country $-i$ that

$$\begin{aligned} \mu_i - \zeta_{-i} + \frac{e_{i,0}}{E_0} &\leq 0 \\ \mu_i - (1-\zeta_i) - \frac{e_{-i,0}}{E_0} &\leq -1 \\ \mu_{-i} - \zeta_i + \frac{e_{-i,0}}{E_0} &\geq 1 > 0. \end{aligned}$$

Hence, if $\bar{E}_i = 0$ for country i, then the FOC (4.2) implies the existence of a unique solution $0 < \bar{E}_{-i} < E_0$ for country $-i$.

To compare $\bar{E}_i$ to $\bar{E}^S$, we need to consider the following three cases.

Case 1. If $0 < \bar{E}_i < \bar{E}_{-i}$, then, by strict concavity of W_i, we obtain the following limits

$$\begin{aligned} \lim_{\bar{E} \to \bar{E}_i} \frac{d\,W(\bar{E})}{d\bar{E}} &= \frac{\mathrm{d}W_{-i}(\bar{E}_i, \hat{\mu}_1)}{\mathrm{d}\bar{E}} > 0, \\ \lim_{\bar{E} \to \bar{E}_{-i}} \frac{d\,W(\bar{E})}{d\bar{E}} &= \frac{\mathrm{d}W_{i}(\bar{E}_{-i}, \hat{\mu}_1)}{\mathrm{d}\bar{E}} < 0. \end{aligned} \tag{A.9}$$

Hence, the INTERMEDIATE VALUE THEOREM implies that $\bar{E}_i < \bar{E}^S < \bar{E}_{-i}$.

Case 2. If $0 = \bar{E}_i < \bar{E}_{-i}$, then Lemma 1 immediately implies that $\bar{E}_i < \bar{E}^S$, while the limit $\bar{E} \to \bar{E}_{-i}$ in (A.9) together with the concavity of W implies that $\bar{E}^S < \bar{E}_{-i}$. Thus, we have that $\bar{E}_i < \bar{E}^S < \bar{E}_{-i}$ again.

Case 3. In the trivial case $\bar{E}_i = \bar{E}_{-i}$, it is obvious that $\bar{E}_i = \bar{E}^S = \bar{E}_{-i}$ holds.

Since $i \in \{1, 2\}$, combining *Case 1 – 3* yields

$$\min\,\{\bar{E}_1,\,\bar{E}_2\} \le \bar{E}^S \le \max\,\{\bar{E}_1,\,\bar{E}_2\},$$

which is the relation stated in Lemma 3. □

Proof of Lemma 4. Again, we need to distinguish three cases.

Case 1. Consider the case where $0 < \bar{E}_i < \bar{E}_{-i}$. Evaluating the derivative of the Nash product at the lower bound of $\mathscr{P}_{\hat{\mu}_1}$, we obtain, by definition of $\bar{E}_i$ and strict concavity of W_{-i}, that

$$\lim_{\bar{E} \to \bar{E}_i} \frac{d\,\mathcal{N}_{\hat{\mu}_1}(\bar{E})}{d\bar{E}} = \frac{\mathrm{d}W_{-i}(\bar{E}_i, \hat{\mu}_1)}{\mathrm{d}\bar{E}} \cdot W_i(\bar{E}_i, \hat{\mu}_1) > 0.$$

By contrast, evaluating the derivative of the Nash product at the upper bound of $\mathscr{P}_{\hat{\mu}_1}$, we obtain, by definition of $\bar{E}_{-i}$ and strict concavity of W_i, that

$$\lim_{\bar{E} \to \bar{E}_{-i}} \frac{d\,\mathcal{N}_{\hat{\mu}_1}(\bar{E})}{d\bar{E}} = \frac{\mathrm{d}W_{i}(\bar{E}_{-i}, \hat{\mu}_1)}{\mathrm{d}\bar{E}} \cdot W_{-i}(\bar{E}_{-i}, \hat{\mu}_1) < 0.$$

The INTERMEDIATE VALUE THEOREM implies that a solution $\bar{E}^B > 0$ to (4.4) in $\mathscr{P}_{\hat{\mu}_1}$ exists. To see that $\bar{E}^B$ is unique and indeed maximizes the Nash product, note first that $W_i > 0$, $\mathrm{d}W_{-i}/\mathrm{d}\bar{E} > 0$ and $\mathrm{d}W_i/\mathrm{d}\bar{E} < 0$ holds for all $\bar{E} \in (\bar{E}_i, \bar{E}_{-i})$. This requires that $W_{-i} > 0$ must hold for any solution $\bar{E}^B$ to (4.4). Hence, for the derivative of the l.h.s. of (4.4) evaluated at $\bar{E}^B$ it holds

$$\begin{aligned} \frac{d^2\,\mathcal{N}_{\hat{\mu}_1}(\bar{E}^B)}{d\bar{E}^2} &= \sum\nolimits_i \left(\frac{\mathrm{d}^2 W_i(\bar{E}^B, \hat{\mu}_1))}{\mathrm{d}\bar{E}^2} \cdot W_{-i}(\bar{E}^B, \hat{\mu}_1)) + \frac{\mathrm{d}W_i(\bar{E}^B, \hat{\mu}_1)}{\mathrm{d}\bar{E}} \cdot \frac{\mathrm{d}W_{-i}(\bar{E}^B, \hat{\mu}_1)}{\mathrm{d}\bar{E}} \right) \\ &< 0. \end{aligned}$$

Graphically, for any solution $\bar{E}^B \in \mathscr{P}_{\hat{\mu}_1}$ to (4.4), the l.h.s. of (4.4) intersects with the r.h.s. (zero) with a negative slope such that only one solution exists, i.e., the solution to (4.4) is unique. Since the second derivative of the Nash product evaluated at $\bar{E}^B$ is negative, $\bar{E}^B$ is indeed a maximizer.

Case 2. Now, consider the case where $0 = \bar{E}_i < \bar{E}_{-i}$. The limit $\bar{E} \to \bar{E}_i$ then admits all signs, i.e., it might be positive or non-positive. In the positive case, $\bar{E}^B > 0$ is determined by the unique solution to (4.4) in $\mathscr{P}_{\hat{\mu}_1}$ following the arguments presented in case 1. In the non-positive case, no solution to (4.4) exists, and $\bar{E}^B = 0$ holds since the Nash product is decreasing on $\mathscr{P}_{\hat{\mu}_1}$.

Case 3. In the trivial case $\bar{E}_i = \bar{E}_{-i}$, it is obvious that $\bar{E}_i = \bar{E}^B = \bar{E}_{-i} > 0$ is the unique solution to (4.4) in $\mathscr{P}_{\hat{\mu}_1}$.

Combining *Case 1 – 3* and recognizing that $\bar{E}^B \leq \max\{\bar{E}_1, \bar{E}_2\} < E_0$ by Lemma 3 yields Lemma 4. □

Proof of Proposition 2. The proof proceeds in three steps. First, we show that

$$\bar{E}^S = \bar{E}^B \qquad \Longleftrightarrow \qquad \hat{\mu}_1 = \mu_1^S, \tag{A.10}$$

denoting μ_1^S the unique allocation that solves $W_i(\bar{E}^S, \hat{\mu}_1) = W_{-i}(\bar{E}^S, \hat{\mu}_1)$, i.e., that equally distributes the welfare level obtained under the efficient cap in both countries. We then establish that $\mu_1^S = 1/2$ in the symmetric case. Continuity thus implies that an allocation $\mu_1^S \in [0, 1]$ also exists if the countries are sufficiently symmetric, i.e., sufficiently similar in terms of benefits, costs, and initial emissions. Third, it is shown that $\hat{\mu}_1 \neq \mu_1^S$ if $B_i(A(\bar{E}^S)) - B_{-i}(A(\bar{E}^S))$ is sufficiently large.

Step 1. We start by establishing that

$$\bar{E}^S = \bar{E}^B \qquad \Longleftrightarrow \qquad W_i(\bar{E}^S, \hat{\mu}_1) = W_{-i}(\bar{E}^S, \hat{\mu}_1), \tag{A.11}$$

Note that $\bar{E}^S$ and $\bar{E}^B$ are unique (cf. Lemma 1 and 4). According to Lemma 1, it holds for the efficient cap that,

$$\sum_i \frac{d\,W_i(\bar{E}^S, \hat{\mu}_1)}{d\bar{E}} = 0. \tag{A.12}$$

Since $\bar{E}^S > 0$ by Lemma 1, we can infer from Lemma 4 that bargaining *can* result in the efficient cap, if and only if the bargaining outcome is determined by the

FOC of the Nash product. In this case, $\bar{E}^B$ satisfies

$$\frac{d\mathcal{N}_{\hat{\mu}_1}(\bar{E})}{d\bar{E}} = \sum\nolimits_i \frac{\mathrm{d}W_i(\bar{E}^B, \hat{\mu}_1)}{\mathrm{d}\bar{E}} \cdot W_{-i}(\bar{E}^B, \hat{\mu}_1) = 0. \tag{A.13}$$

If $\bar{E}^S = \bar{E}^B$, then Equation A.13 implies that

$$\sum\nolimits_i \frac{\mathrm{d}W_i(\bar{E}^S, \hat{\mu}_1)}{\mathrm{d}\bar{E}} \cdot W_{-i}(\bar{E}^S, \hat{\mu}_1) = 0$$

must also hold. Rearranging and inserting (A.12) yields that

$$\frac{\mathrm{d}W_i(\bar{E}^S, \hat{\mu}_1)}{\mathrm{d}\bar{E}} \Big(W_{-i}(\bar{E}^S, \hat{\mu}_1) - W_i(\bar{E}^S, \hat{\mu}_1) \Big) = 0, \qquad i = 1, 2,$$

which implies $W_i(\bar{E}^S, \hat{\mu}_1) = W_{-i}(\bar{E}^S, \hat{\mu}_1)$.

Now, consider the opposite direction. Multiplying both sides of (A.12), we obtain

$$W_i(\bar{E}^S, \hat{\mu}_1)\Big(\frac{d\,W_i(\bar{E}^S, \hat{\mu}_1)}{d\bar{E}} + \frac{d\,W_{-i}(\bar{E}^S, \hat{\mu}_1)}{d\bar{E}} \Big) = 0, \qquad i = 1, 2, \tag{A.14}$$

If $W_i(\bar{E}^S, \hat{\mu}_1) = W_{-i}(\bar{E}^S, \hat{\mu}_1)$, then (A.14) can be rewritten to

$$\sum\nolimits_i \frac{\mathrm{d}W_i(\bar{E}^S, \hat{\mu}_1)}{\mathrm{d}\bar{E}} \cdot W_{-i}(\bar{E}^S, \hat{\mu}_1) = 0 \tag{A.15}$$

Comparing (A.13) to (A.15) immediately yields that $\bar{E}^S = \bar{E}^B$. We can therefore conclude that the relation stated in (A.11) holds.

To establish the following relation

$$W_i(\bar{E}^S, \hat{\mu}_1) = W_{-i}(\bar{E}^S, \hat{\mu}_1) \qquad \Longleftrightarrow \qquad \hat{\mu}_1 = \mu_1^S, \tag{A.16}$$

we need to show that a unique solution μ_1^S to $W_i(\bar{E}^S, \hat{\mu}_1) = W_{-i}(\bar{E}^S, \hat{\mu}_1)$ exists. To do so, note that the following limits obtain

$$\lim_{\hat{\mu}_1 \to \infty} W_1(\bar{E}^S, \hat{\mu}_1) - W_2(\bar{E}^S, \hat{\mu}_1) = +\infty,$$

$$\lim_{\hat{\mu}_1 \to -\infty} W_1(\bar{E}^S, \hat{\mu}_1) - W_2(\bar{E}^S, \hat{\mu}_1) = -\infty,$$

since $\bar{E}^S$ is fixed. Hence, the Intermediate Value Theorem implies the existence of a solution μ_1^S. From calculating the derivatives,

$$\frac{\mathrm{d}W_1(\bar{E}^S, \hat{\mu}_1)}{\mathrm{d}\mu_1} = p(\bar{E}^S)\bar{E}^S > 0, \quad \frac{\mathrm{d}W_2(\bar{E}^S, \hat{\mu}_1)}{\mathrm{d}\mu_1} = -p(\bar{E}^S)\bar{E}^S < 0, \tag{A.17}$$

we can conclude that the allocation μ_1^S is indeed unique. Combining (A.11) and (A.16) directly implies the relation in (A.10).

Step 2. From (2.9), we obtain that $W_i(\bar{E}^S, \hat{\mu}_1) = W_{-i}(\bar{E}^S, \hat{\mu}_1)$ if and only if $\hat{\mu}_1$

solves

$$
\begin{aligned}
&B_i(A(\bar{E}^S)) - B_{-i}(A(\bar{E}^S)) + C_{-i}(a_{-2}(\bar{E}^S)) - C_i(a_i(\bar{E}^S)) = -2x_{i,-i}C_i'(a_i(\bar{E}^S)) \\
&\quad \text{where} \quad x_{i,-i}(\bar{E}) \coloneqq \left(\mu_i \bar{E}^S - e_{i,0} + a_i(\bar{E}^S)\right), \quad \mu_1 = \hat{\mu}_1, \quad \text{and} \quad \mu_2 = 1 - \hat{\mu}_1.
\end{aligned} \tag{A.18}
$$

In the symmetric case, we have have that $B_i(A) = B(A)$, $C_i(a_i) = C(a_i)$ and $e_{i,0} = e_{-i,0}$ for all $i \in \{1, 2\}$. This implies that $a_i(\bar{E}^S) = a_{-i}(\bar{E}^S)$. Hence, the l.h.s. of (A.18) is zero. Inserting $x_{i,-i}$, using that $\bar{E}^S = 2(e_{i,0} - a_i(\bar{E}^S))$ and solving for $\hat{\mu}_1$ yields $\mu_1^S = 1/2$. By continuity, we can now infer that if the countries are sufficiently similar in terms of benefits, costs, and initial emissions, then $\mu_1^S \in [0, 1]$.

Step 3. If $B_i(A(\bar{E}^S)) - B_{-i}(A(\bar{E}^S)) \to \infty$, then the l.h.s. of (A.18) converges to infinity since $C_i(a_i(\bar{E}^S)) \leq C_i(a_i(E_0))$ and $C_{-i}(a_{-i}(\bar{E}^S)) \leq C_{-i}(a_{-i}(E_0))$ are finite. Since all expressions on the r.h.s. of (A.18) besides μ_i are finite, $C_i'(a_i(\bar{E}^S)) \leq C_i'(a_i(E_0))$, $\bar{E}^S \leq E_0$, $a_i(\bar{E}^S) \leq a_i(E_0)$, $e_{i,0} \in \mathbb{R}_{++}$, we must have that $\mu_i \to -\infty$. This implies that

$$
\mu_1^S \to \begin{cases} -\infty & \text{if } i = 1 \\ +\infty & \text{if } i = 2. \end{cases}
$$

Hence, if $B_i(A(\bar{E}^S)) - B_{-i}(A(\bar{E}^S))$ sufficiently large, it holds that $\mu_1^S \notin [0, 1]$ and thus that $\mu_1^S \neq \hat{\mu}_1 \in [0, 1]$. By (A.10), $\mu_1^S \neq \hat{\mu}_1$ directly implies $\bar{E}^S \neq \bar{E}^B$. □

Proof of Corollary 1. Corollary 1 follows from the fact that $\bar{E}^B = \bar{E}^S$ if and only if $\hat{\mu}_1 = \mu_1^S$ by (A.10) and that μ_1^S is, in turn, determined by the unique solution to $W_i(\bar{E}^S, \hat{\mu}_1) = W_{-i}(\bar{E}^S, \hat{\mu}_1)$. Since W_i and W_{-i} also depend on benefit- and cost functions, this allocation generally deviates from an allocation that is solely determined on the basis of initial emissions. □

Proof of Proposition 3. We first establish that

$$
\bar{E}^B = \bar{E}^S \iff \mu_1^B = \mu_1^S \iff \mu_1^S \in [0, 1]. \tag{A.19}
$$

To investigate whether a solution to Problem (5.1) implements $\bar{E}^B = \bar{E}^S$, we solve the relaxed problem

$$
\max_{(\bar{E}, \mu_1)} \mathcal{N}(\bar{E}, \mu_1) \qquad \text{s.t. } 0 \leq \mu_1 \leq 1, \tag{A.20}
$$

and show that the solution to the relaxed problem is in $\mathscr{P}$, i.e., corresponds to a solution to the original Problem (5.1). The Lagrangian of Problem (A.20) writes

$$\mathscr{L}(\bar{E},\mu_1) = W_1(\bar{E},\mu_1)\cdot W_2(\bar{E},\mu_1) - \lambda_1(\mu_1-1) + \lambda_2\mu_1.$$

The solution to the relaxed problem is determined by the system of FOCs

$$\frac{\mathrm{d}\mathscr{L}(\bar{E},\mu_1)}{\mathrm{d}\bar{E}} = \frac{\mathrm{d}W_1(\bar{E},\mu_1)}{\mathrm{d}\bar{E}}\cdot W_2(\bar{E},\mu_1) + W_1(\bar{E},\mu_1)\cdot\frac{\mathrm{d}W_2(\bar{E},\mu_1)}{\mathrm{d}\bar{E}} = 0, \quad \text{(A.21)}$$

$$\frac{\mathrm{d}\mathscr{L}(\bar{E},\mu_1)}{\mathrm{d}\mu_1} = \Big(W_2(\bar{E},\mu_1) - W_1(\bar{E},\mu_1)\Big)p(\bar{E})\bar{E} - \lambda_1 + \lambda_2 = 0, \quad \text{(A.22)}$$

and the complementary slackness conditions

$$\lambda_1(\mu_1-1) = 0, \qquad \text{and} \qquad \lambda_2\mu_1 = 0. \quad \text{(A.23)}$$

Now, three cases can occur.[18]

Case 1. If $\mu_1^S \in [0,1]$, then the tuple $(\bar{E}^S,\mu_1^S)$ is the unique solution to (A.21) and (A.22) by implementing

$$W_1(\bar{E}^S,\mu_1^S) = W_2(\bar{E}^S,\mu_1^S), \qquad \text{and} \qquad \frac{\mathrm{d}W_1(\bar{E}^S,\mu_1^S)}{\mathrm{d}\bar{E}} = \frac{\mathrm{d}W_2(\bar{E}^S,\mu_1^S)}{\mathrm{d}\bar{E}}.$$

To see that $(\bar{E}^S,\mu_1^S) \in \mathscr{P}$, note that $\bar{E}^S$ leads to the greatest level of overall welfare. Hence, for any other cap $\bar{E} \neq \bar{E}^S$, the total level of welfare is lower such that at least one country is worse off, i.e., setting $\bar{E} \neq \bar{E}^S$ does not constitute a Pareto improvement. For $\bar{E} = \bar{E}^S$, on the other hand, any other allocation $\mu_1 \neq \mu_1^S$ leaves one country worse off (cf. (A.17)), i.e., does not constitute a Pareto improvement. Thus, we must have that $(\bar{E}^S,\mu_1^S) \in \mathscr{P}$ such that $(\bar{E}^S,\mu_1^S)$ is also the solution to the original Problem (5.1) in this case.

$(\bar{E}^S,\mu_1^S)$ is indeed the global maximizer of the Nash product, which can be seen as follows. It is well known that the product of two positive real numbers is maximum when the numbers are equal, given that their sum is constant. In our setting, for any given $\bar{E}$, the corresponding level of total welfare is constant, i.e., independent of the actual allocation μ_1. Hence, given $\bar{E}$, the Nash product is maximum for the allocation that equates welfare levels in both countries. Moreover, $\bar{E}^S$ implements the highest positive level of total welfare. The tuple $(\bar{E}^S,\mu_1^S)$ must thus be a global maximizer since it equally distributes the highest level of total welfare.

Case 2. If $\mu_1^S > 1$, then $\bar{E}^S$ does not solve (A.21) since it holds for all $\mu_1 \in [0,1]$

[18] Note that in the case in which no strictly positive solution to (A.21) exists, we have that $0 = \bar{E}^B \neq \bar{E}^S$ by Lemmas 1 and 4.

that

$$\frac{\mathrm{d}W_1(\bar{E}^S,\mu_1)}{\mathrm{d}\bar{E}} = \frac{\mathrm{d}W_2(\bar{E}^S,\mu_1)}{\mathrm{d}\bar{E}}, \quad \text{and} \quad W_1(\bar{E}^S,\mu_1) \neq W_2(\bar{E}^S,\mu_1).$$

Instead, we can read off (A.21) that the solution to the relaxed problem $\bar{E}^R$ must be in the open interval between $\bar{E}_1$ and $\bar{E}_S$. Moreover, (A.21) together with concavity of W_i implies that

$$W_2(\bar{E}^R,\mu_1^R) > W_1(\bar{E}^R,\mu_1^R) \tag{A.24}$$

holds for the solution to the relaxed problem. Hence, from (A.22) we obtain that $\lambda_1 > 0$. By (A.23), we can now infer that $\mu_1^R = 1$ and thus $\lambda_2 = 0$.

To see that the tuple $(\bar{E}^R, 1) \in \mathscr{P}$, first observe that for $\mu_1^R = 1$ any other cap from the open interval between $\bar{E}^R$ and $\bar{E}_2$ leaves country 1 strictly worse off. This is illustrated in Figure 4. Moreover, given any cap from the open interval between $\bar{E}^R$ and $\bar{E}_2$, any other allocation $\mu_1 < 1$ further reduces country 1's welfare. Hence, choosing any cap from the open interval between $\bar{E}^R$ and $\bar{E}_2$ and any $\mu_1 \in [0,1]$ does not constitute a Pareto improvement. Second, since $\bar{E}^R$ is in the open interval between $\bar{E}_S$ and $\bar{E}_1$, any cap in the open interval between $\bar{E}_1$ and $\bar{E}^R$ leads to a lower level of overall welfare irrespective of the allocation μ_1, see Figure 4. Hence, for any allocation $\mu_1 \in [0,1]$, any cap from this interval leaves at least one country worse off, i.e., does not constitute a Pareto improvement. The tuple $(\bar{E}^R, 1)$ is thus Pareto-efficient, that is $(\bar{E}^R, 1) \in \mathscr{P}$. The solution to the relaxed Problem (A.20) therefore corresponds to the solution to the original Problem (5.1) such that $\bar{E}^B = \bar{E}^R \neq \bar{E}^S$ and $\mu_1^B = \mu_1^R = 1 < \mu_1^S$.[19]

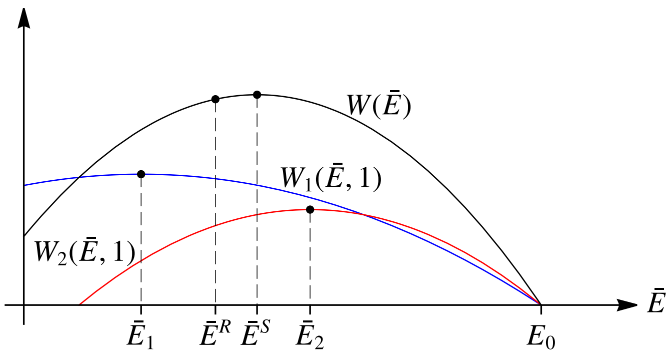

Figure 4: *Pareto efficiency of the solution to the relaxed problem.*

Case 3. If $\mu_1^S < 0$, then $\bar{E}^B = \bar{E}^R \neq \bar{E}^S$ and $\mu_1^B = \mu_1^R = 0 > \mu_1^S$ follows by

[19] It can readily be verified that $(\bar{E}^R, 1)$ is indeed the maximizer of the relaxed problem.

analogous arguments.

Combining these three cases yields the relation stated in (A.19).

Part (i) of Proposition 3 now immediately follows from (A.19) and the fact that $\mu_1^S \in [0,1]$ if the countries are sufficiently symmetric, which we have shown in the proof of Proposition 2.

Part (ii) of Proposition 3 results from (A.19) and the fact that $\mu_1^S \notin [0,1]$ if the countries' benefits are sufficiently different, which we have already argued in the proof of Proposition 2. □

Proof of Proposition 4. We know from Proposition 3 that, in the absence of outside options, bargaining implements $(\bar{E}^S, \mu_1^S)$ where $\mu_1^S \in [0,1]$ if the countries are sufficiently symmetric. Hence, it remains to show that the tuple $(\bar{E}^S, \mu_1^S)$ also satisfies the additional constraints $W_1(\bar{E}, \mu_1) \geq W_1^N$ and $W_2(\bar{E}, \mu_1) \geq W_2^N$ for sufficiently symmetric countries. To do so, we show for the (completely) symmetric case where $B_i(A) = B(A)$, $C_i(a_i) = C(a_i)$ and $e_{i,0} = e_{-i,0}$ that $W_i(\bar{E}^S, \mu_1^S) > W_i^N$ holds for all $i \in \{1,2\}$. Continuity thus implies that, as long as the countries are sufficiently symmetric, $(\bar{E}^S, \mu_1^S)$ satisfies the additional constraints, i.e., $(\bar{E}^S, \mu_1^S)$ is indeed the solution to Problem (5.2).

In the symmetric case without outside options, the bargaining outcome is $(\bar{E}^S, 1/2)$, which leads to the following welfare levels:

$$W_i(\bar{E}^S, 1/2) = W_{-i}(\bar{E}^S, 1/2) = \frac{W^S}{2}. \tag{A.25}$$

For national caps, on the other hand, symmetry leads to $\bar{e}_i^N = \bar{e}_{-i}^N$ and thus

$$W_i^N = W_{-i}^N = \frac{W^N}{2}. \tag{A.26}$$

In view of Proposition 1, combining (A.25) and (A.26) immediately implies that $W_i(\bar{E}^S, \mu_1^S) > W_i^N$ holds for all $i \in \{1,2\}$ in the symmetric case. □

Proof of Proposition 5. Observing that bargaining over the amount and allocation of permits implements $\bar{E}^S$ if and only if $\mu_1^S \in [0,1]$ by (A.19), the proposition follows immediately from comparing Examples 2 and 3. □