Working Papers

www.cesifo.org/wp

Anticipated International Environmental Agreements

Ömer T. Açıkgöz
Hassan Benchekroun

CESifo Working Paper No. 6194
Category 9: Resource and Environment Economics
November 2016

ISSN 2364-1428

CESifo
Center for Economic Studies & Ifo Institute

CESifo *Working Paper No. 6194*

Anticipated International Environmental Agreements

Abstract

Consider a situation in which countries anticipate an international environmental agreement (IEA) to be in effect sometime in the future. What is the impact of the future IEA on current emissions after its announcement? We show that the answer to this question is ambiguous. We examine four types of IEAs that aim to reduce pollution stock in the environment. IEA type 1 sets a level of emissions, IEA type 2 and IEA type 3 set a percentage and a uniform cut vis-à-vis the business-as-usual policy respectively. IEA type 4 sets the policy that maximizes future joint benefits of the signatories. We show that all of these agreements reach their goal in the long run, but the intended benefits of these IEAs can potentially be offset by the anticipatory non-cooperative response of the signatories, leading to an environmental degradation in the short run. Which IEA is preferable depends on the targeted level emissions during the phase of cooperation. When this target is above a certain threshold, welfare under IEA type 1 is larger than under IEA type 2 which is larger than under IEA type 3. Moreover, the highest level of welfare that can be attained under IEA type 1 is above the highest level of welfare achieved by any one of the other three IEAs.

JEL-Codes: Q530, Q540, Q580, Q590.

Keywords: international environmental agreements, climate agreement, future agreements, transboundary pollution, dynamic games.

Ömer T. Açıkgöz
Department of Economics
Yeshiva University
USA - New York, NY 10033
acikgoz@yu.edu

Hassan Benchekroun
Department of Economics & CIREQ
McGill University
Canada - H3A-2T7 Montreal QC
hassan.benchekroun@mcgill.ca

We thank the Associate Editor, Peter Berck, three anonymous referees, Michèle Breton, Robert Cairns, Bard Harstad, Ngo Van Long and Cees Withagen for helpful comments. Usual disclaimers apply. Hassan Benchekroun thanks the Canadian Social Sciences and Humanities Research Council (SSHRC) and the Fonds de recherche du Québec – Société et culture (FRQSC) for financial support.

1 Introduction

Transboundary pollution problems give rise to prisoners' dilemma type of situations that are particularly difficult to resolve due to the absence of a supra-national authority that can impose an allocation of pollution rights, or enforce agreed upon contracts. Any agreement between the parties involved needs to be self-enforcing. The difficulty of reaching an agreement can in principle be overcome by an ad hoc system of transfers and incentives to prevent free-riding.[1] In practice, the number of international treaties to protect the environment or resources has been steadily increasing since the second half of the twentieth century. The International Environmental Agreements (IEA) Database Project lists environmental agreements from 1800 to 2013.[2] In the project, agreements are defined as environmental if "they seek, as a primary purpose, to manage or prevent human impacts on natural resources; plant and animal species (including in agriculture, since agriculture modifies both); the atmosphere; oceans; rivers; lakes; terrestrial habitats; and other elements of the natural world that provide ecosystem services.". The project lists over 1200 multilateral environmental agreements, more than half of which were signed after 1990, 1500 bilateral environmental agreements, and 250 "other" environmental agreements. The "Other" category includes environmental agreements between governments and international organizations or non-state actors. The list includes conventions, treaties, protocols, amendments, modifications.

The process of defining precise commitments for each member of an agreement, the ratification and the passing of domestic legislation to enforce the agreement's commitments is often very tedious and lengthy. An interesting feature of many such agreements is that there is a substantial period of time that separates the date of adoption and the date of entry into force.[3,4] For example, in the case of climate change, the United Nations Framework Convention on Climate Change (UNFCCC) was signed by 154 countries in the Earth Summit in Rio de Janeiro in 1992 and entered into force in 1994. However, binding limits on greenhouse gases emissions, in the case of developed countries, were only defined in 1997 under the Kyoto Protocol; a protocol that entered into force in 2005.[5,6] Even bilateral

[1] See e.g., Germain, Toint, Tulkens and de Zeeuw (2003) or Petrosyan and Zaccour (2003) for the case of stock pollution game and Calvo and Rubio (2013) for a recent survey.

[2] http://iea.uoregon.edu/page.php?file=home.htm&query=static
Version 2013.2, July 2013.

[3] Barrett (1998) breaks down the process of treaty making into five stages: pre-negotiation, negotiation, ratification, implementation and renegotiation. Congleton (2001) identifies four phases of IEAs. The first stage is the recognition of the benefits from cooperation and gives rise to "Symbolic Treaties,"in the second stage parties sign a "Procedural Treaty" defining procedures to evaluate alternative policy targets, in the third stage parties agree on specific targets and sign a "Substantive" treaty and in the fourth "Domestic" stage each party passes legislation to conform with its treaty obligations.

[4] Barrett (2003), in the appendix of Chapter 6, gives a list of multilateral environmental agreements and provides the date of adoption and the date of ratification.

[5] Canada withdrew in 2011 and the U.S. did not ratify the treaty.

[6] Other examples of delayed entry into force of environmental treaties include The Convention for the Protection of the Marine Environment of the North-East Atlantic which was opened for signature in 1992 and entered into force in 1998. The Convention on Long-Range Transboundary Air Pollution, CLRTAP, open for signature in 1979 and entered

environmental treaties can take decades to be finalized. (See Congleton (2001) for examples)

In situations where the damage is caused by the accumulation of pollution, the anticipation of cooperation in the future may influence countries' emission levels over the pre-cooperation phase. In this paper, we investigate two related issues on the anticipated IEAs: (i) "How are the pre-cooperation emission levels affected by the announcement of an IEA?", and (ii) "How do these anticipation effects influence the design of IEAs that feature implementation lags?". To this end, we consider a transboundary stock pollution game in which countries will adopt an emission policy with the intention of reducing pollution stock compared to business-as-usual (BAU) policy. For the sake of simplicity, we abstract from modeling the political process by which a cooperative solution is reached. In this environment, a number of countries produce homogeneous products, and pollution is emitted as a by-product. Emissions of pollution accumulate, constituting a transboundary stock of pollution, damaging all countries in the same way.

To provide a benchmark case, following the literature, we first determine the equilibrium under the BAU scenario, where countries behave non-cooperatively over the two periods. This is contrasted with the first-best cooperative equilibrium where countries follow a profile of emissions strategies for each period that maximizes the aggregate welfare. It is well known that *conditional on a given stock of pollution*, a Markov-Perfect Nash Equilibrium (MPNE) of the game results in larger emissions of pollution than the socially efficient level, i.e. where each country adopts the pollution strategy that maximizes the joint welfare of all countries.[7] We prove this result within the context of our model, and reveal the possibility that equilibrium level of emissions under cooperation can actually *exceed* those under BAU *in the future*. Although this possibility arises rarely (with extreme parameter values), at face value, it indicates that the goal of cutting emissions in the future and the goal of achieving a higher welfare might be conflicting. On the other hand, we prove that first-best solution unambiguously leads to smaller pollution stock both in the short run and in the long run vis-à-vis BAU. Given these findings, and taking the first-best solution as an appealing benchmark, the comparison of pollution stocks in the short run, i.e. the stock at the time the IEA enters into force, and the long run, i.e., the stock at the end of the game, is more informative about the performance of an IEA compared to BAU.

Our first thought experiment considers a case in which countries adopt an IEA in period two that is fully anticipated in period one. We examine the impact of four different types of IEAs on the emission level over the pre-cooperation phase. Given that the first-best solution is more environmentally friendly than the solution of the BAU scenario both in the short run, and in the long run, it is natural to seek an IEA with the same qualities. We illustrate, within the confines of these four IEAs we chose for their

into force in 1983, Marpol 73/78 ("Marpol" is short for marine pollution and 73/78 short for the years 1973 and 1978) the International Convention for the Prevention of Pollution From Ships was signed in 1973. The current convention entered into force in 1983.

Even the Vienna Convention on the Law of Treaties (VCLT), a treaty on the international law on treaties between states was adopted in 1969 and entered into force in 1980.

[7] See the surveys of dynamic pollutions games in Jorgensen, Martin-Herran and Zaccour (2010) or Long (2011).

subjective appeal, that this is a very demanding property. We prove that *all* these IEAs reach their goals in the long run, i.e. they ultimately lead to a pollution stock lower than that under BAU. However, we illustrate, for each IEA, how the intended benefits can potentially be curtailed by the anticipatory non-cooperative response of the signatories, leading to an environmental degradation in the short run, relative to BAU.

In the first type of IEA, countries anticipate that the level of emissions in period 2 will be set at an agreed upon target, independent of the pollution stock, assuming that the countries can commit to this level. In the other three IEAs, countries cannot commit to an emission *level*, but rather commit to an emission *policy rule* that depends on the level of the pollution stock. In the second and third types of IEAs, the anticipated emission policy represents a percentage, and a constant cut of future emissions vis-à-vis BAU respectively. The fourth type of IEA, which serves as a benchmark for our exercise, represents the ex post first-best emission policy, i.e. the policy which maximizes the joint welfare of all countries (in period 2).

In the first type of anticipated IEA, IEA type 1, where countries can commit to a target level of emissions, we show that a cap in period 2, set exactly at the BAU emissions level, results in a decrease of current emissions with respect to the BAU emissions level. However, emissions in period 1 is a decreasing function of the target level for period 2: A tighter target in period 2 results in larger emissions in period 1. Therefore, the overall impact of a cap on period 2 emissions is ambiguous.

In the case of an anticipated IEA type 2, we show that a marginal cut in emissions leads to an increase of current emissions if and only if the elasticity of marginal utility with respect to the emission level is larger than one. We then illustrate that when the elasticity is large enough, a sufficiently large cut in emissions leads to a *decrease* in emissions in period 1. A similar ambiguity result is illustrated in the linear-quadratic framework: We show that the anticipation of an IEA results in a decrease of current emissions if the damage from pollution is sufficiently large. Otherwise, the impact of an IEA is ambiguous, and typically depends on the level of the stock of pollution.

We show that for IEA type 3, a small cut in period 2 emissions increases pre-cooperation emissions, however a large enough cut will decrease them. For the fourth type of IEA, IEA type 4, where countries adopt the ex post first-best emission policy, we give a necessary condition under which the future IEA results in a decrease of current emissions with respect to their BAU level. We show that this necessary condition is never satisfied in the linear-quadratic case, and in the case where utility function is logarithmic.

It is widely recognized that, even if cooperation is expected in the future, the actual policies ratified by domestic legislatures will diverge from policies prescribed under an 'ideal' scenario in period 2, (IEA type 4 in our context), policies that maximize the joint welfare of all countries in period 2. The reality of policy making is that the final legislation that will be implemented is the outcome of an interaction of several interest groups.[8] In this sense, we view IEAs of type 1, 2 and 3 to be potentially more relevant

[8] See Oates and Portney (2003) or Wangler, Altamirano-Cabrera and Weikard (2011).

for policy considerations. The larger the extent of the departure from the BAU outcome, the larger the resistance that the policy makers will meet. The influence of interest groups, whether environmental groups or groups representing industrial interests, on environmental policy is well-established. (See e.g., Oates and Portney (2003) or Wangler et al. (2011) for a survey) The important difference between IEA type 1 and IEAs type 2 and 3 is the level of the BAU outcome that is taken as a benchmark with respect to which the abatements are made. IEAs of type 2 and 3 represent a scenario where lobbyist can still influence the abatement pledge of their country in an IEA up to the moment the IEA enters into force. In these cases, the IEA's goal has to be conditioned on the prevailing stock of pollution at the time the IEA enters into force in order for the agreement to be subgame perfect. By contrast, IEA type 1 captures a situation where lobbying occurs only in the negotiation phase and therefore countries have the ability to commit to a future level of emissions even before the IEA enters into force. IEA type 4, where countries are able to maximize their joint welfare in period 2 and do not cooperate in period 1, represents a hybrid scenario between the scenario where countries cooperate over the two periods, and the scenario, where non-cooperation prevails in both periods.

Following the analysis of the anticipation effects of various IEAs on emissions levels prior to the cooperative phase, we examine, given a target level of emissions for period 2, which IEA yields the highest level of welfare. When period 2's emission target is above a certain threshold, we show that IEA type 1 yields a higher welfare than IEA type 2 and IEA type 3 that yield the targetted period 2's emission level. If players cannot commit to an emission level for period 2, when period 2's emission target is above a certain threshold, an IEA type 2 is preferable to an IEA type 3. We also examine for each type of IEA which level of period 2 emissions yields the largest welfare. Our analysis also reveals that the highest level of welfare achieved under an IEA type 1 is above the highest level of welfare achieved by any one of the other three IEAs. We also show that welfare under IEA type 4 (ex post first-best announcement), is a *lower bound* for welfare for the other types. From an implementation point of view, this implies that IEA type 4 is unambiguously dominated by IEAs type 1, 2 and 3.

Next, we step outside the boundaries of the four types of IEAs and focus on the case where the planner is not constrained to be within a specific IEA class, aside from mild technical restrictions to ensure tractability. We prove that even in this general environment, *no* IEA can implement first-best level of emissions and welfare, however, suitable design of an IEA that takes into account the anticipation effects can get *arbitrarily* close to them. Our proof is constructive and provides a guideline on the design and implementation of IEAs with a time lag. This optimistic result is unfortunately overshadowed by the increased complexity of such suitably designed IEAs, as the number of periods before the entry into force of an agreement increases; a complexity that limits their policy relevance.

1.1 Literature Review

The possibility that an environmentally-friendly policy may end up having negative environmental consequences, as is the case under all four types of IEAs considered, is by now well-established: see e.g., Hoel (1991 and 1992) or Eichner and Pethig (2011) in the case where environmentally friendly policies are unilaterally adopted by a sub-group of countries or Benchekroun and Ray Chaudhuri (2014) in the case where countries adopt a cleaner technology.[9]

A related literature that highlights the role of fossil fuels in the climate change problem examines the role of environmentally friendly policies on the intertemporal extraction of fossil fuels. Within this framework, the unintended negative impact of environmentally-friendly policies (such as subsidy to cleaner energies) on the environment was coined 'the green paradox' by Sinn (2008) who examines the reaction of the supply side of fossil fuels to the implementation of "greening (demand-reducing) policies" (such as a tax on fossil fuels) by importing countries. It is shown that fossil fuel exporters may respond by accelerating the extraction of their stock, and thereby accelerating the accumulation of pollution, worsening the damage caused by fossil fuels. Both in Sinn (2008) and the related literature (see e.g., Sinn (2012), Van der Ploeg, and Withagen (2012), Grafton, Kompas and Long (2012) or for surveys, Van der Werf and Di Maria (2012), Van der Ploeg and Withagen (2013) or Long (2014)) the source of a green paradox is the reallocation of the supply of fossil fuels that result from a greening policy.

Few papers explicitly examine the impact of announcement of a future environmental policy on current emissions. Smulders, Tsur and Zemel (2012) study the impact of announcing a carbon tax to be implemented in the future. They show that this announcement results in an increase of carbon emissions in the period that precedes the implementation of the tax, compared to the BAU scenario. The result is driven by an increase in savings and the accumulation of capital, which in turn translates into more production and emissions. Di Maria, Smulders and Van der Werf (2012) examine the effectiveness of environmental policy in the presence of an implementation lag within the confines of a non-renewable resources model. They show that an implementation lag can result in an increase of emissions during the period that precedes implementation. This result is driven by the scarcity of the available resources. In our model, it is the countries' strategic response to the anticipation of an IEA that potentially leads the countries' cutting back (or increasing) current emissions.

Countries' strategic response to environmental policy is also at the root of the findings of Eichner and Pethig (2011). In a two period framework, they consider a world made of three types of countries: fossil fuel owners, fossil users that have an active environmental policy, and fossil fuel users that have no environmental policy. They show that the tightening of the environmental policy of the abaters in period 1 can result in an increase of that period's emissions of the non-abaters (carbon leakage) which may even overcompensate the decrease of abaters' emissions and result in an increase of total emissions during that period (green paradox). These results are shown in a fairly general setup and within a general equilibrium framework. Analogous results are shown in the case where the environmental policy takes effect in period 2, as is the case of IEAs in our paper. In our paper, all countries are abaters in period 2 and the focus is on the intertemporal leakage only. A key difference with Eichner and Pethig (2011)

[9] See Benchekroun (2003) in the context of international fisheries.

is that they consider a problem where the total level of emissions over period 1 and 2 is fixed, and an environmental policy can only shift the allocation of emissions across the two periods. In our paper the total level of emissions over the two periods is endogenous and an IEA would not only affect the allocation of emissions across the two periods but also affect the amount of pollution emitted over the two periods.[10] Moreover, our paper also examines the role played by the type of IEA envisioned for period 2 and offers insights into the design of a future IEA.

Harstad (2015) considers a model of formation of consecutive IEAs. Countries are assumed to be able to contract over emissions; however the length of the contract is finite. Along with the choice of emissions, each country chooses a level of investment in a technology that reduces its 'need for pollution'. When choosing its investment and emission levels, each country anticipates that there will be future negotiations. It is shown that IEAs may result in a smaller welfare than under the BAU scenario, where no IEA is signed. This is because when countries anticipate future negotiations, the hold-up problem diminishes the incentives to invest in technology. This negative outcome is more likely to arise, the weaker the intellectual property rights, the shorter the length of the agreement, and the larger the number of countries. In contrast with Harstad (2015), where IEAs start at time zero, we have an explicit pre-agreement period where emissions are decided non-cooperatively, and we focus on emission decisions only. Moreover, in Harstad (2015) and Harstad (2012) the equilibrium in the BAU case is not affected by the IEA that will prevail in the future. This feature is due to (i) the possibility to invest in technology and (ii) the fact that the investment costs are linear.[11]

Our paper is closely related to that of Beccherle and Tirole (2011), who also examine the impact of the announcement of a future environmental agreement within a two-period game between two countries. In their framework, the damage from pollution is generated in period 2, and depends only on the environmental policies adopted in period 2. However, in period 1, each country can adopt an environmental policy that will impact its private utility in period 2. The paper examines the impact of delaying an agreement on the bargaining outcome between the two countries, focusing on the environmental policies adopted in period 2. While their public good framework is quite general, a crucial assumption of their model is that a more lax environmental policy in period 1 results in a more lax environmental policy in period 2. Such an assumption makes their analysis not applicable to the case of a transboundary stock pollution game, where the emission policies in period 1 and period 2 are typically negatively related. Indeed, in our model, an increase in emissions in period 1 increase the stock

[10] As recent developments in the fossil fuels markets have shown, the reserves of non-conventional oil and gas are virtually limitless, much like the reserves of coal, and such fuels become increasingly economically exploitable as technolgy improves and the energy price increases.

[11] Our result also contrasts with Harstad (2014) which shows in the case of a game between the owner of a fossil fuel deposit and a player seeking to prevent the use of fossil fuels, that the prospect of future compensation for conservation (in period 2) leads to more present conservation (in period 1). The contrast is due to the nature of the goods considered, i.e., "conservation goods" defined as goods where the buyer does not consume the good but only seeks "to prevent the seller from consuming it in the future."

of pollution in period 2, resulting in a decrease in period 2 emissions. Another contrast with Beccherle and Tirole (2011) is that we do not explicitly model the bargaining taking place in period 2. Rather, we assume that in period 2, the emission policy is fixed through an agreement and is independent of emissions in period 1.[12]

In the following section, we present the model and analyze the BAU and first-best benchmarks. In section 3, we examine the impact of four different types of anticipated future agreements on current emissions. Section 4 includes a welfare comparison of the four IEAs, section 5 takes a mechanism design approach and discusses the ex ante optimal choice of IEA. Section 6 offers concluding remarks.

2 Model

There are two periods $t \in \{1, 2\}$, and n identical countries $i \in I \equiv \{1, ..., n\}$. We assume that there is a linear technology that results in emissions as a by-product of consumption goods production. Let P_t denote the initial stock of pollution at time t.

The sequence of events in each period is as follows:

1. Given an initial stock of pollution P_t, countries choose their emissions $(E_{t,1}, .., E_{t,n})$ to produce consumption goods. Countries derive utility from consumption, and due to linear technology, we assume, without loss of generality, that utility takes the form $u(E_{t,i})$.
2. The sum of the current emissions and the inherited stock cause damage $D\Bigg(P_t + \sum_{k=1}^{n} E_{t,k}\Bigg)$.
3. Natural decay of pollution (current emissions plus the initial stock of pollution) occurs. We have

$$P_2 = \left(P_1 + \sum_{k=1}^{n} E_{1,k}\right)(1 - \delta)$$

   where $\delta \in [0, 1]$ represents the natural rate of decay of the stock of pollution.

We assume that the discount factor, $\beta \in [0, 1]$, is identical for all n countries. The utility and damage functions are fixed over time, identical for all countries, and are of common knowledge. Relaxing these assumptions would not qualitatively alter our results. We make the following technical assumptions on the utility and damage functions:

Assumption 1. *Utility and damage functions are twice continuously differentiable and satisfy*

1. $u'(E) > 0$ *for all* $E > 0$*,* $D'(0) = 0$*, and* $D'(P) > 0$ *for all* $P > 0$*.*

2. $u''(E) < 0$ *for all* $E > 0$*, and* $D''(P) > 0$ *for all* $P > 0$*.*

[12]However, in the case of IEAs of type 2, 3 and 4 in our game, the outcome of the agreement in period 2 still depends on decisions made in period 1 since those decisions impact the stock of pollution in period 2 and therefore impact the outcome of the agreement.

3. $\lim_{E\to\infty} u(E) - D(P+E) = -\infty$ *for all* $P \geq 0$.

The first two parts of assumption 1 are standard and ensure that optimization problem faced by agents is smooth and convex. For most results, either one of the strict inequalities in part 2 can be replaced with a weak one. The third part ensures that damage component overwhelms the benefit of emissions asymptotically, and is used to establish existence of solutions when the choice set for emissions is not compact. Note that this last assumption is in conformity with the view that there is a threshold level of pollution concentration beyond which ecological catastrophic regime shifts occur, and habitats are irreversibly destroyed. While there can be disagreements over the precise level of such threshold its existence is widely accepted. We impose additive separability of utility and damage function for ease of exposition. This feature is *not* essential for any of our results provided that we impose obvious cross-derivative restrictions to preserve strict concavity of the objective functions. From this point onwards, we will assume that assumption 1 holds for all of our results, unless we explicitly state otherwise.

For some of our results, we need to make the following additional technical assumption to ensure the existence and uniqueness of a Nash equilibrium. This assumption is not very exclusive and covers most functions considered in the literature. In particular, it is satisfied by linear-quadratic utility/damage functions, constant-relative-risk-aversion (CRRA) and constant-absolute-risk-aversion (CARA) utility functions.

Assumption 2. *Utility and damage functions are thrice continuously differentiable and satisfy* $u''' \geq 0$ *and* $D''' \leq 0$.

In the non-cooperative setup, each country i chooses an emission strategy, taking the profile of emission strategies of the other countries as given. We seek a Markov-Perfect Nash Equilibrium of this game and solve this problem backwards, starting from the strategies in period 2. After we characterize the properties of the non-cooperative problem, we provide a comparison with the first-best solution, where the planner chooses a vector of emissions in each period that maximizes the joint discounted sum of welfare.

2.1 Period 2

Given a stock of pollution P_2, under the non-cooperative scenario, country i takes $E_{2,-i} \geq 0$, the vector of emissions of all the countries except country i as given, and chooses its emissions $E_{2,i}$ as a solution to the following problem:

$$\max_{E\geq 0} u(E) - D\left(P_2 + E + \sum_{k\in I\setminus\{i\}} E_{2,k}\right) \tag{1}$$

Since countries are identical, all equilibria are symmetric (to be proven below) and we drop the country indices. The equilibrium strategy, $E_{2,NC}$, is a function of the only state variable P_2, the pollution stock inherited from period 1. Next, we characterize its properties.

Proposition 1. *Under assumption 1, there exists a unique Nash equilibrium for period 2. The equilibrium is symmetric, and the strategy* $E_{2,NC} : \mathbb{R}_+ \to \mathbb{R}_+$ *solves*

$$u'\left(E_{2,NC}(P_2)\right) \leq D'\big(P_2 + nE_{2,NC}(P_2)\big) \text{ with equality if } E_{2,NC}(P_2) > 0. \tag{2}$$

In addition, $E_{2,NC}$ *has the following properties:*

1. *It is continuously differentiable and* $E'_{2,NC}(P_2) \in [-\frac{1}{n}, 0]$ *for all* $P_2 > 0$*, except at* $\bar{P}_2 \equiv \lim_{E\downarrow 0}(D'^{-1}\big(u'(E)\big)$ *if this point is well-defined.*
2. *If, in addition, assumption 2 holds, equilibrium strategy* $E_{2,NC}(P_2)$ *is convex, and the value of following this strategy,* $V_2(P_2) \equiv u(E_{2,NC}(P_2)) - D(P_2 + nE_{2,NC})$*, is decreasing and concave in* P_2*.*

Proof: See Appendix A.

Remark: We would like to point out that even though $E_{2,NC}(P_2)$ is a decreasing function of P_2, the induced final pollution stock $P_3 \equiv P_2 + nE_{2,NC}(P_2)$ is *increasing* in P_2. This follows from $E'_{2,NC}(P_2) \in [-\frac{1}{n}, 0]$.

For future reference we provide an expression for the derivative of the strategy at an interior solution. Total differentiation of expression (2) with respect to P_2 gives

$$u''\left(E_{2,NC}\left(P_2\right)\right)E'_{2,NC}\left(P_2\right) - \left(1 + nE'_{2,NC}\left(P_2\right)\right)D''\left(P_2 + nE_{2,NC}\left(P_2\right)\right) = 0 \tag{3}$$

which implies

$$E'_{2,NC}\left(P_2\right) = \frac{D''\left(P_2 + nE_{2,NC}\left(P_2\right)\right)}{u''\left(E_{2,NC}\left(P_2\right)\right) - nD''\left(P_2 + nE_{2,NC}\left(P_2\right)\right)} \in [-\frac{1}{n}, 0]. \tag{4}$$

2.2 Period 1

We now consider the problem of country $i \in I$ in period 1 under the non-cooperative scenario. Given an initial stock of pollution P_1, country $i \in I$ takes the Nash equilibrium policy in period 2, $E_{2,NC}\left(P_2\right)$ as well as the vector of emissions in period 1 of all the countries except for country i, $E_{1,-i}$, as given and solves the following problem

$$\max_{E\geq 0} u\left(E\right) - D\big(P_1 + E + \sum_{k\in I\setminus\{i\}} E_{1,k}\big) + \beta V_2(P_2) \tag{5}$$

where $V_2(P_2)$ is the value for period 2 as defined in proposition 1 and

$$P_2 = \big(P_1 + E + \sum_{k\in I\setminus\{i\}} E_{1,k}\big)\left(1-\delta\right).$$

Next, we characterize the properties of the Nash equilibrium strategy $E_{1,NC}(P_1)$.

Proposition 2. *Suppose assumptions 1 and 2 hold. Then, there exists a unique Nash equilibrium for period 1. The equilibrium is symmetric, and the strategy $E_{1,NC} : R_+ \to R_+$ satisfies*

$$u'\left(E_{1,NC}\left(P_1\right)\right) \leq D'\left(P_1 + nE_{1,NC}\left(P_1\right)\right) + \beta\left(1-\delta\right)\left(1+\left(n-1\right)E'_{2,NC}\right)u'\left(E_{2,NC}\left(P_2\right)\right) \tag{6}$$

for all P_1, with equality if $E_{1,NC}(P_1) > 0$.

Proof: The objective in problem (5), $f(E)$, is strictly concave under assumption 1 and the properties of $V_2(.)$ established in proposition 1. Moreover, $-f(E)$ is coercive over the convex choice set $E \geq 0$.[13] Therefore there exists a unique solution E that satisfies the Kuhn-Tucker necessary condition

$$\begin{aligned} u'\left(E\right) \leq D'\left(P_1 + E + \sum_{k \in I \setminus \{i\}} E_{1,k}\right) \\ - \beta \frac{dP_2}{dE}\left[u'\left(E_{2,NC}\left(P_2\right)\right)E'_{2,NC}\left(P_2\right) - \left(\left(1+nE'_{2,NC}\left(P_2\right)\right)\right]D'\left(P_2 + nE_{2,NC}\left(P_2\right)\right)\right) \end{aligned}$$

for all P_1, with strict equality when $E > 0$. Using the facts that $E_{2,NC}$ satisfies equation (2) and that $\frac{dP_2}{dE} = 1 - \delta$ to simplify this expression further, it is straightforward to show that $E_{1,NC}\left(P_1\right)$ is a symmetric equilibrium if and only if condition (6) holds. Given the convexity of $E_{2,NC}(P_2)$ already established in proposition 1, proof of existence and uniqueness of a solution $E_{1,NC}(P_1)$ to equation (6) follows analogous steps to the ones used in the proof of proposition 1. We omit the details. Since marginal damage is common for all countries (even if they were to emit different amounts), and utility is strictly concave, there cannot be a non-symmetric equilibrium. ■

We will refer to the emission policies $E_{1,NC}\left(P_1\right)$ and $E_{2,NC}\left(P_2\right)$ as the business-as-usual (BAU) policies in period 1 and 2 respectively. Condition (6) represents the usual balance of the marginal benefit from current emissions with the present and future marginal damages from pollution. We would like to draw the attention of the reader to the term $\left(n-1\right)E'_{2,NC}$. The presence of this term captures the impact of a given country's current emissions on the future emissions of other countries. In period 1, while any given country takes the strategies of the other countries as given, if the strategies are stock-dependent, by affecting the stock of pollution, a country can manipulate the emission of other countries. In the dynamic games literature, this channel of interaction is referred to as the feedback effect. This term $\left(n-1\right)E'_{2,NC}$ is negative as shown in proposition 1, therefore the feedback effect reduces the future marginal damage from pollution and boosts the incentive of each country to increase its current emissions compared to a situation where this channel is absent (e.g., if countries choose emission paths instead of emission policies).

[13] A function $g : \mathbb{R}^n \to \mathbb{R}$ is coercive if $g(x) \to \infty$ as $||x|| \to \infty$. This holds in our problem due to item 3 in assumption 1.

2.3 First-Best Solution

Before we examine the impact of anticipation of an IEA, we characterize some properties of the first-best solution and how it compares to BAU. The objective of the unconstrained planner choosing $(E_1, E_2) \geq 0$ can be written as

$$f(E_1, E_2; P_1) \equiv u(E_1) - D(P_1 + nE_1) + \beta\Big(u(E_2) - D\big((1-\delta)(P_1 + nE_1) + nE_2\big)\Big) \tag{7}$$

The planner's problem is

$$W_{coop}(P_1) \equiv \max_{E_1, E_2 \geq 0} f(E_1, E_2; P_1) \tag{8}$$

Proposition 3. *Under assumption 1, there exists a unique solution* $(E_{1,Coop}, E_{2,Coop})$ *to planner's problem* (8) *for all* $P_1 \geq 0$*, and it solves*

$$u'(E_{2,Coop}) \leq nD'(P_2 + nE_{2,Coop}) \text{ with equality if } E_{2,Coop} > 0 \tag{9}$$
$$u'(E_{1,Coop}) \leq n\Big(D'\big(P_1 + nE_{1,Coop}\big) + \beta(1-\delta)D'(P_2 + nE_{2,Coop})\Big) \text{ with equality if } E_{1,Coop} > 0,$$

where $P_2 \equiv (1-\delta)(P_1 + nE_{1,Coop})$.

Proof: By assumption 1, objective $f(.)$ is *strictly* concave and $-f(.)$ is coercive. The constraint set $E_1, E_2 \geq 0$ is non-empty, closed and convex. Hence, there exists a unique solution characterized by Kuhn-Tucker first-order necessary conditions given in the proposition.■

Remark: The planner's solution $(E_{1,Coop}, E_{2,Coop})$ established above is contingent on the given initial pollution level P_1, but it is straightforward to express the second period emissions in Markovian form (contingent on P_2) by solving this problem by backward induction. Clearly this solution $E_{2,Coop}(P_2)$ would also satisfy conditions (9) and we omit the proof.

It is well-known that under non-cooperation, free-riding prevails and in our framework it can be shown that we have

$$E_{2,Coop}(P_2) < E_{2,NC}(P_2) \text{ and } E_{1,Coop}(P_1) < E_{1,NC}(P_1)$$

for all $P_1, P_2 \geq 0$ and such that $E_{1,NC}(P_1) > 0$ and $E_{2,NC}(P_2) > 0$. When P_1, P_2 are such that $E_{1,NC}(P_1) = 0$ and $E_{2,NC}(P_2) = 0$, we have $E_{1,Coop}(P_1) = 0$ and $E_{2,Coop}(P_2) = 0$.

Interestingly, even though the cooperative emission policy is never above the BAU emission policy, it does not imply that cooperative period 2 emission level is never above the period 2 emission level under BAU. Under the first-best scenario, emissions in period 1, and therefore the stock of pollution at the start of period 2, can be so small that the second period emissions could be *larger* than those under BAU. This case arises when agents care little about smoothing emissions over time (close to linear utility) and the damage function has a very high curvature (for instance, a cubic function). The intuition can be summarized as follows: When marginal damage is very high, decay of pollution(δ) is low, and utility is close to linear, efficiency requires postponing emissions as much as possible, because agents suffer the consequences of first period marginal emissions over both periods. Under BAU, the private marginal damage is dampened due to the "feedback effect" and agents pollute more in comparison. Although this effect is always present, when utility is linear, the feedback effect is maximized.[14] As a result, under cooperation, agents enter the second period with a much lower pollution stock compared to BAU. When the pollution stock and the curvature of the marginal damage are large enough, second period emissions under BAU are below their first-best levels.

We use the example of a linear utility and cubic damage function to illustrate this possibility.[15] More specifically suppose $u'(E) = 1$, and $D'(P) = P^2$. The same result holds under CRRA utility with a low value of CRRA coefficient σ, and any damage function that has higher curvature than a quadratic function. Using the first-order necessary conditions in propositions 1, 2 and 3, the policy functions (ignoring corner solutions) satisfy

$$E_{1,Coop}(P_1) = \left(\frac{1}{n}\right)^{3/2} \sqrt{1-\beta(1-\delta)} - \frac{P_1}{n} \qquad E_{2,Coop}(P_2) = \left(\frac{1}{n}\right)^{3/2} - \frac{P_2}{n}$$

$$E_{1,NC}(P_1) = \frac{1}{n}\sqrt{1 - \frac{\beta(1-\delta)}{n}} - \frac{P_1}{n} \qquad E_{2,NC}(P_2) = \frac{1}{n} - \frac{P_2}{n}$$

Consider the case where $\beta = 1$, $\delta = 0$ and $P_1 = 0$. It can be verified that the path of emissions would be $E_{1,Coop} = 0$, $E_{1,NC} = \frac{1}{n}\sqrt{\frac{n-1}{n}}$, $E_{2,Coop} = \left(\frac{1}{n}\right)^{3/2}$, $E_{2,NC} = \frac{1}{n}\left(1-\sqrt{\frac{n-1}{n}}\right)$. Clearly, $E_{2,Coop} > E_{2,NC}$ for all $n \geq 2$.

While we cannot rely on a clear ranking between the cooperative and non-cooperative level of emissions in period 2, we show that, at the end of each period, the stock of pollution under non-cooperation is never below the stock of pollution under cooperation. More precisely, we will refer to equilibrium levels $P_{2,Coop}$ and $P_{2,NC}$ as the pollution stock in the short run under first-best and BAU respectively. Analogously, we will call equilibrium levels $P_{3,Coop} \equiv (1-\delta)(P_1+nE_{1,Coop})+nE_{2,Coop}$ and $P_{3,NC} \equiv (1-\delta)(P_1+nE_{1,NC})+nE_{2,NC}$, the long-run pollution stock under first-best and under BAU respectively. The next proposition establishes a robust feature of the first-best solution which we will take as a reference to compare against IEAs: The cooperative solution is environmentally conservative, *both in the short run and in the long run.*

Proposition 4. *Under assumption 1,*

1. *Long run equilibrium pollution stock satisfies* $P_{3,Coop} \leq P_{3,NC}$ *for all* $P_1 \geq 0$.

2. *Short run equilibrium pollution stock satisfies* $P_{2,Coop} \leq P_{2,NC}$ *for all* $P_1 \geq 0$.

Proof: See Appendix A.

[14] Maximized in the sense that the policy function has the highest slope achievable in absolute value, $-\frac{1}{n}$.

[15] This specification violates assumptions 1 and 2 since $u'' = 0$ and $D''' \geq 0$, however, it can be verified by going over the steps of the proofs that all conclusions in propositions 1, 2 and 3 are still valid under this particular specification. Observe that the equilibrium strategy is linear (therefore convex), and all problems faced by agents are convex. Also, there can be non-symmetric equilibria, but we ignore them here as their presence does not change our main point.

3 Anticipated IEAs

Having provided the benchmark case, we now examine the impact of a fully anticipated IEA, under which all n countries would implement an emission strategy in period 2, denoted by $E_{2,C}(P_2)$ and that is aimed at reducing pollution in the environment relative to BAU. In the previous section, we have seen that first-best solution might feature second period emissions higher than BAU. On the other hand, first-best solution is generically environmentally friendly both in the short run and in the long run, independent of preference and damage functions. (Proposition 4) Given this property, from a welfare point of view, it is more natural to seek IEAs that share the latter feature. For this reason we will use the following terminology and criterion to judge whether an IEA is environmentally friendly relative to BAU.

Definition 1. *Given an initial pollution stock $P_1 \geq 0$, an IEA $E_{2,C}(P_2)$ is*

1. ***Environmentally friendly in the short run** (relative to BAU) if the equilibrium pollution stock satisfies $P_{2,C} \leq P_{2,NC}$,*

2. ***Environmentally friendly in the long run** (relative to BAU) if the equilibrium pollution stock satisfies $P_{3,C} \leq P_{3,NC}$.*[16]

In the remainder of the paper we consider four different types of IEAs be implemented in period 2 but anticipated by all the countries at the beginning of period 1: (i) $E_{2,C}(P_2) = \bar{E}$ for all $P_2 \geq 0$, (ii) $E_{2,C}(P_2) = (1-\varepsilon) E_{2,NC}(P_2)$ for all $P_2 \geq 0$, (iii) $E_{2,C}(P_2) = E_{2,NC}(P_2) - \eta$ for all $P_2 \geq 0$ and (iv) $E_{2,C}(P_2) = E_{2,Coop}(P_2)$ for all $P_2 \geq 0$. These IEAs will be referred to as IEA type 1, type 2, type 3, and type 4 respectively.

An IEA type 1 prescribes a specific target of the level of emissions in period 2. Under this type of IEA, countries are assumed to be able to commit to a level of emissions in period 2, even if the stock of pollution in period 2 turns out to be different from the level anticipated. This arrangement is similar to the commitment assumed in open-loop games and results in equilibria that may not be subgame perfect.

IEAs of type 2 and 3 correspond to a cut, percentage in case of 2 and uniform in case of 3, with respect to the BAU policy. Note that they represent a cut with respect to an emission *policy* and not a level of emissions (as in case 1). Using emission policies for period 2 yields a more robust IEA than an IEA type 1, since the agreement is stock dependent. If the stock of pollution in period 2 differs from what was originally anticipated, the principle of the agreement is still valid. This robustness is similar to that of a subgame perfect equilibrium in a non-cooperative framework.[17] The IEAs of type 2 and 3 are inspired from a negotiation process where the benchmark is the BAU scenario, and where parties

[16] As we defined in the previous section, $P_{3,C}$ and $P_{3,NC}$ are the final (long-run) equilibrium pollution stocks under the given IEA and under BAU respectively.

[17] See Dockner et al. (2000) Ch 4.

try to find improvements over the BAU policies. The size of the departure, captured by either ε or η, reflects the limitations of policy makers to impose changes with respect to a BAU scenario because of political pressure and lobbying from different groups in the economy.

The IEA type 4 is provided as a benchmark. It represents the 'ideal' agreement that can be reached in period 2. Under an IEA type 4 policy makers are assumed free of any domestic political constraints, and can choose the policy that maximizes global welfare. While it is a useful benchmark, in particular in the assessment of the potential gains from an IEA, or the absence of an IEA, it has proven to be too difficult to reach in practice. Most of the IEAs take the form of a reduction with respect to a BAU scenario.

We separately examine the impact of the four IEAs in the short run, i.e., on P_2 (or on first period equilibrium emissions) and in the long run, i.e., on P_3.

3.1 Short-run Impact

We define the short-run impact of an anticipated IEA as the impact of an IEA on the stock of pollution at the start of period 2, P_2. Since there is a one-to-one relationship between the equilibrium level of emissions in period 1 and P_2, the short-run impact of an IEA can be gauged by looking at the impact on the first period emissions directly.

Given an announcement $E_{2,C}(P_2)$, and period 1 strategies of all other countries, $E_{1,-i}$, a country i solves the following problem in period 1:

$$\max_{E\geq 0} u\left(E\right) - D\Bigg(P_1 + E + \sum_{k\in I\setminus\{i\}} E_{1,k}\Bigg) + \beta\left[u\left(E_{2,C}(P_2)\right) - D\big(P_2 + nE_{2,C}(P_2)\big)\right] \tag{10}$$

where

$$P_2 = \Bigg(P_1 + E + \sum_{k\in I\setminus\{i\}} E_{1,k}\Bigg)\left(1-\delta\right).$$

Assuming that $E_{2,C}(.)$ is differentiable where it is positive, we obtain the following characterization of the first period non-cooperative response $E_{1,C}(P_1)$ using the first-order necessary condition for problem (10) and imposing symmetry:

$$u'(E_{1,C}) \leq D'(P_1+nE_{1,C})+\beta(1-\delta)\Big\{[1+(n-1)E'_{2,C}(P_{2,C})]D'(P_{3,C})-E'_{2,C}(P_{2,C})\big(u'(E_{2,C}(P_{2,C}))-D'(P_{3,C})\big)\Big\} \tag{11}$$

with equality if $E_{1,C} > 0$, where $P_{2,C} = (1-\delta)(P_1 + nE_{1,C})$ and $P_{3,C} = P_{2,C} + nE_{2,C}(P_{2,C})$.

3.1.1 IEA Type 1: Setting an emission target

We examine the case in which

$$E_{2,C}\left(P_2\right) = \bar{E} \text{ where } 0 \leq \bar{E} < E_{2,NC}\left(P_{2,NC}\right),$$

and where $P_{2,NC}$ is the equilibrium stock of pollution at the beginning of period 2 in the non-cooperative scenario. Thus $E_{2,NC}(P_{2,NC})$ represents the equilibrium emissions in period 2 in the absence of an IEA in period 2. Since $\bar{E} < E_{2,NC}(P_{2,NC})$, at face value, this IEA is aimed at providing a cleaner environment relative to BAU. Whether it actually is environmentally friendly in the sense we defined it in the previous section depends critically on the countries' anticipatory response in period 1. We investigate this next.

How do countries react to this announcement in period 1? Assuming a symmetric equilibrium, the first-order necessary condition (11) implies an emission strategy $E_{1,C}(P_1)$ that solves

$$u'(E_{1,C}) \leq D'(P_1 + nE_{1,C}) + \beta(1-\delta) D'\left((P_1 + nE_{1,C})(1-\delta) + n\bar{E}\right) \text{ with equality if } E_{1,C} > 0 \tag{12}$$

This problem is well-defined and has a unique solution under assumption 1. We omit the details since the steps of the proof are virtually identical to those in propositions 1 and 2.

Lemma 1. *For an IEA type 1, a decrease in the threshold $\bar{E}$ unambiguously results in an increase in current emissions and therefore an increase in $P_{2,C}$.*

Proof: See Appendix B.1.

We should note, however, that setting $\bar{E} = E_{2,NC}(P_{2,NC})$ does not yield $E_{1,C} = E_{1,NC}$. Indeed, when $\bar{E} = E_{2,NC}(P_{2,NC})$, we have $E_{1,C} < E_{1,NC}$, a result proven in the lemma below. This is due to the feedback effect that is present in the non-cooperative case (i.e., $E'_{2,NC}(P_2) < 0$).[18]

Lemma 2. *An IEA type 1 that features $\bar{E} = E_{2,NC}(P_{2,NC})$ is environmentally friendly in the short run.*

Proof: See Appendix B.1.

Thus the mere announcement of an IEA that will set an emission target in period 2 equal to the BAU emission level in period 2, results in a downward shift of period 1 emissions and therefore of the initial stock of pollution in period 2, P_2. We can infer from Lemmas 1 and 2 that the overall impact of setting emissions in period 2 to a level $\bar{E} < E_{2,NC}(P_{2,NC})$ on current emissions is ambiguous. It clearly results in a decrease of current emissions if $\bar{E}$ is sufficiently close to $E_{2,NC}(P_{2,NC})$, whereas setting $\bar{E}$ at level close to zero may result in an increase of current emissions.[19] This result is summarized in the following proposition:

Proposition 5. *The impact of an IEA type 1 on the equilibrium emissions in period 1 is ambiguous, therefore it may not be environmentally friendly in the short run.*

[18]The feedback effect of production constraints has been highlighted in Dockner and Haug (1990, 1991) in the context of dynamic import quotas.

[19]See appendix B.1 for a proof in the case where the damage from pollution is such that $D'(0) = 0$.

3.1.2 IEA Type 2: A percentage cut in emissions

Suppose now the countries anticipate an IEA of the following type:

$$E_{2,C}(P_2) = (1-\varepsilon) E_{2,NC}(P_2) \text{ where } \varepsilon \in [0,1].$$

The intent of this IEA is clear when $\varepsilon < 1$. The policy targets a cleaner environment by prescribing lower emissions in period 2 relative to BAU, conditional on the second period pollution stock, i.e. $E_{2,C}(P_2) < E_{2,NC}(P_2)$ for all P_2. We investigate the first period response to the announcement of the IEA. Assuming a symmetric equilibrium and an interior solution, the first-order necessary condition (11) implies an emission strategy $E_{1,C}(P_1)$ that solves

$$u'(E_{1,C}(P_1)) - D'(P_1 + nE_{1,C}(P_1)) = R(\varepsilon) \tag{13}$$

where

$$R(\varepsilon) \equiv -\beta(1-\delta)\Lambda$$

with

$$\Lambda \equiv u'((1-\varepsilon)E_{2,NC}(P_2))(1-\varepsilon)E'_{2,NC}(P_2) - \left[1 + n(1-\varepsilon)E'_{2,NC}(P_2)\right] D'(P_2 + n(1-\varepsilon)E_{2,NC}(P_2)).$$

The following technical lemma is useful in identifying under which conditions IEA type 2 is environmentally friendly in the short run:

Lemma 3. *We have*

$$E_{1,NC}(P_1) > E_{1,C}(P_1) \text{ for all } P_1$$

if and only if

$$R(0) < R(\varepsilon).$$

Proof: This follows from the facts that (i)

$$u'(E_{1,C}(P_1)) - D'(P_1 + nE_{1,C}(P_1)) = R(\varepsilon)$$
$$u'(E_{1,NC}(P_1)) - D'(P_1 + nE_{1,NC}(P_1)) = R(0)$$

and, (ii) $u'(X) - D'(P_1 + nX)$ is a strictly decreasing function of X.■

To establish the sign of $R(0) - R(\varepsilon)$, we examine the impact of a change in ε on $R(.)$. For tractability reasons, we first examine the impact of a marginal proportional reduction of emissions with respect to the BAU scenario, i.e., marginal change in ε in the neighborhood of $\varepsilon = 0$.

Proposition 6. *A type 2 IEA that features a marginal proportional reduction in emissions is environmentally friendly in the short run if and only if* $\sigma \equiv -\frac{E_{2,NC}(P_2)u''(E_{2,NC}(P_2))}{u'(E_{2,NC}(P_2))} < 1$:

Proof: See Appendix B.2. This result (and in fact almost all of our results) extend to an infinite horizon setting. To make this point concrete, Appendix B.2 also contains an extension of the proof of proposition 6 to infinite horizon version of the model.

In light of Proposition 6, it is immediate that in the case where utility is of CRRA form, i.e.,

$$u(E) = \frac{E^{1-\sigma}}{1-\sigma}$$

we have $R'(0) > 0$ or $E_{1,NC}(P_1) > E_{1,C}(P_1)$ if and only if $\sigma < 1$.

We argue that this result does not carry over to large $\varepsilon > 0$. It is easy to construct an example where small ε leads to an *increase* in current emissions, and large ε leads to a *decrease* in current emissions relative to BAU levels. Suppose utility function is CRRA type with $\sigma > 1$. Fix P_2, and consider the limit of $R(\varepsilon; P_2)$ as $\varepsilon \to 1$. Under CRRA utility, $E_{2,NC}(P_2) > 0$ holds for all $P_2 \geq 0$, and since $\sigma > 1$, $\lim_{\varepsilon\to 1} u'((1-\varepsilon)E_{2,NC}(P_2))(1-\varepsilon) = \infty$. Under CRRA, $E'_{2,NC}(P_2) < 0$ holds and parts of $R(\varepsilon; P_2)$ that involve $D'(.)$ remain bounded in the limit. Therefore $R(\varepsilon; P_2) \to \infty$ for all $P_2 \geq 0$ as $\varepsilon \to 1$. The first-order condition (13) implies $E_{1,C} \to 0$ as a response. To sum up, by choosing ε sufficiently high, one can make first period emissions $E_{1,C}$ arbitrarily small, in particular, it can be made smaller than BAU level $E_{1,NC}$. Proposition 6 suggests that small ε leads to an increase in current emissions relative to BAU when $\sigma > 1$. These two results combined make it clear that the impact of IEA type 2 on first period emissions is ambiguous in general. CRRA utility with $\sigma > 1$ is an important case also from an empirical point of view. There is no consensus on the value of the CRRA coefficient, especially in the macroeconomics literature, largely due to the peculiar feature that the inverse of the coefficient represents the elasticity of intertemporal substitution (EIS). Estimates on CRRA coefficient vary, but most studies find empirical support for a value larger than one, and an EIS value smaller than one whenever it is the subject of study. (See Friend and Blume (1975), Hall (1988), Holt and Laury (2002), Havranek et al. (2015) among others.)

Along with the specification in which the utility exhibits constant elasticity, another widely used framework is the linear-quadratic (LQ) model, where utility is a quadratic function of emissions, and the damage function is a quadratic function of the stock of pollution. We show below that the sign of $R'(0)$ is ambiguous even in the LQ case:

$$u(E) = \left(A - \frac{B}{2}E\right)E$$

and

$$D(P) = \frac{1}{2}\gamma P^2.$$

For the rest of the exposition, we impose $A = 1$ and $B = 1$ without loss of generality. Note that $u(E)$ achieves its maximum at $E = 1$. It can be shown, at a symmetric equilibrium, that the emissions strategies satisfy

$$E_{2,NC}(P_2) = \max\left\{0, \frac{1-\gamma P_2}{1+n\gamma}\right\}$$

and

$$E_{1,C}\left(P_1\right) = \max\left\{0, \frac{\Gamma + \Phi P_1}{\Delta}\right\}$$

where

$$\Delta = 1 + [3 + \beta(1-\delta)^2]n\gamma + [3n + \beta(1-\delta)^2(1+\varepsilon^2 + 2\varepsilon\left(n-1\right))]n\gamma^2 + [1 + \beta(1-\delta)^2\varepsilon^2]n^3\gamma^3$$

$$\Gamma = 1 + \gamma\left[2n - \beta\left(\delta - 1\right)\left(-1+\varepsilon\right)\left(\varepsilon + n\right)\right] + \gamma^2 n\left[n - \beta\left(\delta-1\right)\left(\varepsilon - 1\right)\left(1 + n\varepsilon\right)\right],$$

$$\Phi = -\gamma\left\{1 + \beta\left(1-\delta\right)^2 + \left(1 + \beta\left(1-\delta\right)^2\varepsilon^2\right)\gamma^2 n^2 + \gamma\left(\beta\left(1-\delta\right)^2\left(1+\varepsilon^2 + 2\varepsilon\left(n-1\right)\right) + 2n\right)\right\}.$$

In the special case where $\varepsilon = 0$, we have $E_{1,C}(P_1) = E_{1,NC}(P_1)$. Moreover, substitution of u'' and u' gives

$$\sigma = \frac{\left(E_{2,NC}\left(P_2\right)\right)^2}{1 - E_{2,NC}\left(P_2\right)}$$

The sign of $\sigma - 1$ is difficult to determine, since the expression of P_2 is non-trivial. We can however state that in the limit case where γ tends to zero, one would expect $E_{2,NC}\left(P_2\right)$ to be close to 1, since $u\left(E\right)$ achieves its maximum at $E = 1$ and the damage from pollution tends to zero. Therefore when γ tends to zero σ tends to infinity, implying that a marginal proportional reduction of future emissions results in an increase of current emissions. In another limit case where γ becomes arbitrarily large, one could expect $E_{2,NC}\left(P_2\right)$ to tend to zero and therefore σ would tend to zero, implying that a marginal proportional reduction of future emissions results in a decrease of current emissions. We provide more precise statements for the analysis of the sign of $\frac{dE_{1,C}(P)}{d\varepsilon}$ below. For ease of exposition, we examine the cases $P_1 = 0$ and $P_1 > 0$ separately.

The case $P_1 = 0$:

It can be shown that

$$\left.\frac{dE_{1,C}(0)}{d\varepsilon}\right|_{\varepsilon=0} = \frac{\beta\left(1-\delta\right)\gamma\left(n-1\right)\left(1+n\gamma\right)\Omega}{\Delta^2}$$

where

$$\Omega = 1 + \left(-1 + \beta\left(1-\delta\right)^2 + 2\delta\right)\gamma n - n^3\gamma^3 + \gamma^2 n\left(\beta\left(1-\delta\right)^2 + n\left(2\delta - 3\right)\right).$$

Therefore the sign of $\left.\frac{dE_{1,C}(0)}{d\varepsilon}\right|_{\varepsilon=0}$ is the same as that of Ω.

The term Ω is a cubic function of γ. Its graph in a (γ, Ω) space has an inverted N shape and there exist γ_1 and $\gamma_2 > \gamma_1$ where Ω reaches a local minimum and maximum respectively. The function is an increasing function of γ over (γ_1, γ_2) and decreasing elsewhere.

Lemma 4. *There exists a unique $\bar{\gamma} > 0$ such that $\Omega < 0$ for $\gamma > \bar{\gamma}$ and $\Omega > 0$ for $0 < \gamma < \bar{\gamma}$.*

Proof: See Appendix B.2.

We now summarize our main result for this case in the proposition below.

Proposition 7. *For $P_1 = 0$, there exists a unique $\bar{\gamma} > 0$ such that for $\gamma > (<)\bar{\gamma}$, an IEA type 2 that features a marginal proportional cut in emissions by all players is environmentally (not) friendly in the short run.*

Proof: This follows from the fact that the sign of $\left.\frac{dE_{1,C}(0)}{d\varepsilon}\right|_{\varepsilon=0}$ is the same as the sign of Ω, and from Lemma 4.■

Note that Ω can also be rewritten as

$$\Omega = \left(n\gamma\left(\delta-1\right)^2 + n\gamma^2\left(\delta-1\right)^2\right)\left(\beta-\bar{\beta}\right)$$

where

$$\bar{\beta} \equiv \frac{n^3\gamma^3 + \left(3-2\delta\right)n^2\gamma^2 - \left(2\delta-1\right)n\gamma - 1}{n\gamma\left(\delta-1\right)^2 + n\gamma^2\left(\delta-1\right)^2}.$$

Thus Ω is an affine function of β with a positive slope, since $\left(n\gamma\left(\delta-1\right)^2 + n\gamma^2\left(\delta-1\right)^2\right) > 0$ and a sign of the intercept that is ambiguous. When $n\gamma$ is small enough we have $\bar{\beta} < 0$ and therefore $\left.\frac{dE_{1,C}(0)}{d\varepsilon}\right|_{\varepsilon=0} > 0$. When $n\gamma$ is large enough $\bar{\beta} > 1$ and therefore for any $\beta \in [0,1]$ we have $\Omega < 0$ and therefore $\left.\frac{dE_{1,C}(0)}{d\varepsilon}\right|_{\varepsilon=0} < 0$. For 'intermediate' values of $n\gamma$, we have $\bar{\beta} \in (0,1)$ and then the sign of Ω will depend on the value that β takes: for $\beta \in (\bar{\beta}, 1]$ we have $\Omega > 0$ and $\left.\frac{dE_{1,C}(0)}{d\varepsilon}\right|_{\varepsilon=0} > 0$ whereas for $\beta \in (0, \bar{\beta})$ we have $\left.\frac{dE_{1,C}(0)}{d\varepsilon}\right|_{\varepsilon=0} < 0$. This is intuitive, when the discount factor plays a role, players that are patient enough react to the IEA by reducing their emissions in period 1.

We can also rewrite Ω as

$$\Omega = n\beta\gamma\left(1+\gamma\right)\delta^2 - 2n\gamma\left(\gamma\left(\beta-n\right) + 2\left(\beta-1\right)\right)\delta + \left(n\gamma^2\left(\beta-3n\right) - n^3\gamma^3 + n\gamma\left(\beta-1\right) + 1\right).$$

This is a quadratic U-shaped function of δ. We have

$$\left.\frac{d\Omega}{d\delta}\right|_{\delta=0} = -2n\gamma\left(\gamma\left(\beta-n\right) + 2\left(\beta-1\right)\right) > 0 \text{ since } \beta \in [0,1] \text{ and } n \geq 2$$

and therefore $\frac{d\Omega}{d\delta} > 0$ for all $\delta \geq 0$. Thus if $n\gamma$ is small enough we have $\Omega|_{\delta=0} = \left(n\gamma^2\left(\beta-3n\right) - n^3\gamma^3 + n\gamma\left(\beta-1\right) + 1\right) > 0$ and therefore $\Omega > 0$ and $\left.\frac{dE_{1,C}(0)}{d\varepsilon}\right|_{\varepsilon=0} > 0$ for all $\delta > 0$. If $n\gamma$ is large enough we have $\Omega|_{\delta=0} = \left(n\gamma^2\left(\beta-3n\right) - n^3\gamma^3 + n\gamma\left(\beta-1\right) + 1\right) < 0$ and therefore the sign of Ω depends on δ. Then there exists a threshold $\bar{\delta} > 0$ such that for $\delta \in (0, \bar{\delta})$ we have $\Omega < 0$ and thus $\left.\frac{dE_{1,C}(0)}{d\varepsilon}\right|_{\varepsilon=0} < 0$. When $\delta > \bar{\delta}$ we have $\left.\frac{dE_{1,C}(0)}{d\varepsilon}\right|_{\varepsilon=0} > 0$.

The case $P_1 > 0$:

Let $F(\varepsilon) \equiv \frac{\Phi}{\Delta}$. After substitution and derivation with respect to ε we obtain

$$F(\varepsilon) = -\frac{\gamma\left\{1 + \beta\left(1-\delta\right)^2 + \left(1 + \beta\left(1-\delta\right)^2\varepsilon^2\right)\gamma^2 n^2 + \gamma\left(\beta\left(1-\delta\right)^2\left(1+\varepsilon^2+2\varepsilon\left(n-1\right)\right) + 2n\right)\right\}}{1 + [3+\beta(1-\delta)^2]n\gamma + [3n+\beta(1-\delta)^2(1+\varepsilon^2+2\varepsilon\left(n-1\right))]n\gamma^2 + [1+\beta(1-\delta)^2\varepsilon^2]n^3\gamma^3}$$

for which it can be shown that

$$F'(0) = -\frac{2(n-1)\beta\gamma^2(\delta-1)^2(n\gamma+1)^2}{(3n^2\gamma^2 + n^3\gamma^3 + 3n\gamma + n\beta\gamma + n\beta\gamma^2 + n\beta\gamma\delta^2 - 2n\beta\gamma^2\delta + n\beta\gamma^2\delta^2 - 2n\beta\gamma\delta + 1)^2} < 0.$$

When emissions are positive, an increase in ε results in a decrease of the slope, a steeper emission strategy. These expressions imply the following result.

Proposition 8. *There exists a unique $\bar{\gamma} > 0$ such that for $\gamma > \bar{\gamma}$, an IEA type 2 that features a marginal proportional cut in emissions by all players is environmentally friendly in the short run for all $P_1 \geq 0$.*

For $\gamma \leq \bar{\gamma}$, the impact of such an IEA depends on initial pollution stock P_1 : There exists $\bar{P} > 0$ such that it is environmentally friendly in the short run if and only if $P_1 \geq \bar{P}$.

3.1.3 IEA Type 3: A constant cut in emissions

In this section, we examine the case where countries anticipate an IEA in period 2 that features a constant cut in emissions:[20]

$$E_{2,C}(P_2) = E_{2,NC}(P_2) - \eta \text{ where } \eta \in [0, E_{2,NC}(P_{2,NC})].$$

Assuming a symmetric equilibrium, and an interior solution, the first-order necessary condition (11) implies an emission strategy $E_{1,C}(P_1)$ that solves

$$u'(E_{1,C}(P_1)) - D'(P_1 + nE_{1,C}(P_1)) = Q(\eta)$$

where

$$Q(\eta) \equiv -\beta(1-\delta)\left(u'(E_{2,NC}(P_2) - \eta)E'_{2,NC}(P_2) - \left(1 + nE'_{2,NC}(P_2)\right)D'(P_1 + nE_{2,NC}(P_2) - n\eta)\right)$$

Comparing $E_{1,C}(P_1)$ with $E_{1,NC}(P_1)$ amounts to comparing $Q(\eta)$ and $Q(0)$. Since $u''(E_{1,NC}(P_1)) - D''(P_1 + nE_{1,NC}(P_1)) < 0$, we have $E_{1,C}(P_1) < E_{1,NC}(P_1)$ if and only if $Q(\eta) > Q(0)$. The sign of $Q(\eta) - Q(0)$ is difficult to determine in general. As in the case of a proportional emission reduction, we evaluate the sign of $Q(\eta) - Q(0)$ in the case of a marginal decrease of emissions with respect to the BAU policy.

Proposition 9. *An IEA type 3 that features a marginal cut in emissions by all players is not environmentally friendly in the short run.*

Proof: See Appendix B.3.

In the LQ model, it can be shown that the result of proposition 9 extends to the case of non-marginal constant cut of emissions. However there exist functional forms of utility functions for which

[20]We should clarify that we actually assume that $E_{2,C}(P_2) = \max\{0, E_{2,NC}(P_2) - \eta\}$, but we will not make this explicit for notational simplicity.

this extension does not hold: for some initial pollution levels $P_1 \geq 0$ and for some $\eta > 0$, first period emissions can *decrease* relative to BAU level of emissions. Indeed, consider the following simple example: Take any CRRA type utility function $u(.)$, which we know satisfies $\lim_{E \downarrow 0} u'(E) = \infty$. Fix initial pollution $P_1 \geq 0$ and let $\eta \equiv E_{2,NC}(P_{2,NC})$ be the level of emissions in period 2 that would prevail under BAU. If IEA takes the form $E_{2,C}(P_2) = E_{2,NC}(P_2) - \eta$, when all countries emit BAU levels (or higher) in period 1, they would not be allowed to pollute in period 2. Obviously this cannot be a Nash equilibrium: Polluting marginally less in period 1 is infinitely rewarding since doing so will lead to a slightly lower P_2 and positive emissions (in fact, for all agents) in period 2.[21] We conclude that the impact of a type 3 IEA on first period emissions is ambiguous. Small values of η unambiguously *increases* current emissions due to proposition 9, but as our example illustrates, sufficiently large η may *decrease* current emissions relative to BAU.

3.1.4 IEA Type 4: Ex post first-best emissions

We examine the case where countries anticipate that there will be full cooperation in period 2: The chosen vector of emission strategies will be the one that maximizes the joint welfare of all the countries in period 2. i.e. $E_{2,C}(P_2) = E_{2,Coop}(P_2)$. We call this agreement, the ex post first-best IEA. In period 1, each country acts non-cooperatively, choosing its emission strategy that maximizes its private welfare. We show that the anticipation effects might lead to environmental degradation in the short run. At a symmetric equilibrium, period 1 strategy, denoted by $E_{1,C}(P_1)$ solves

$$\begin{aligned} & u'(E_{1,C}(P_1)) - D'(P_1 + nE_{1,C}(P_1)) \\ & + \beta(1-\delta)\left\{u'(E_{2,Coop}(P_2))E'_{2,Coop}(P_2) - \left[1 + nE'_{2,Coop}(P_2)\right]D'(P_2 + nE_{2,Coop}(P_2))\right\} \\ = \;& 0 \end{aligned}$$

In proposition 3, we have shown that under full cooperation, in period 2 we have $u'(E_{2,Coop}(P_2)) - nD'(P_2 + nE_{2,Coop}(P_2)) = 0$, and therefore the expression can be simplified further to

$$u'(E_{1,C}(P_1)) - D'(P_1 + nE_{1,C}(P_1)) - \beta(1-\delta)D'(P_2 + nE_{2,Coop}(P_2)) = 0.$$

In the following proposition, we provide a necessary condition for this particular IEA to result in a decrease of current emissions.

Proposition 10. *If IEA type 4 is environmentally friendly in the short run, then the following condition must hold:*

$$\frac{u'(E_{2,Coop}(P_{2,Coop}))}{u'(E_{2,NC}(P_{2,NC}))} - n\left(1 + (n-1)E'_{2,NC}(P_{2,NC})\right) > 0. \tag{14}$$

Proof: See Appendix B.4.

[21] By proposition 1, $E'_{2,NC} \leq 0$ and this inequality is strict under CRRA assumption.

We argue below that the necessary condition for $E_{1,NC}(P_1) > E_{1,C}(P_1)$ given in the above proposition is never satisfied in the linear-quadratic (LQ) framework and the case where utility function is logarithmic.

First of all, if $E_{1,NC}(P_1) > E_{1,C}(P_1)$ were satisfied, then we have $P_{2,NC} > P_{2,Coop}$. Ignoring the possibility of optimal zero emissions, since cooperative emissions strategy is downward sloping, we have

$$E_{2,Coop}(P_{2,NC}) < E_{2,Coop}(P_{2,Coop})$$

or

$$u'(E_{2,Coop}(P_{2,NC})) > u'(E_{2,Coop}(P_{2,Coop}))$$

This implies

$$\frac{u'(E_{2,Coop}(P_{2,Coop}))}{u'(E_{2,NC}(P_{2,NC}))} - n\left(1 + (n-1)E'_{2,NC}(P_{2,NC})\right) < \frac{u'(E_{2,Coop}(P_{2,NC}))}{u'(E_{2,NC}(P_{2,NC}))} - n\left(1 + (n-1)E'_{2,NC}(P_{2,NC})\right)$$

However, we show below in the LQ and Log cases that

$$\frac{u'(E_{2,Coop}(P_2))}{u'(E_{2,NC}(P_2))} - n\left(1 + (n-1)E'_{2,NC}(P_2)\right) < 0 \text{ for all } P_2,$$

and in particular

$$\frac{u'(E_{2,Coop}(P_{2,NC}))}{u'(E_{2,NC}(P_{2,NC}))} - n\left(1 + (n-1)E'_{2,NC}(P_{2,NC})\right) < 0.$$

Therefore, the necessary condition for $E_{1,NC}(P_1) > E_{1,C}(P_1)$ in proposition 10 is never satisfied. It must then be the case that $E_{1,NC}(P_1) \leq E_{1,C}(P_1)$.

Linear-Quadratic Case: When $u(E) = E\left(1 - \frac{1}{2}E\right)$ and $D(P) = \frac{1}{2}\gamma P^2$, after substitution of the optimal strategies, it can be shown that

$$\frac{u'(E_{2,Coop}(P_2))}{u'(E_{2,NC}(P_2))} - n\left(1 + (n-1)E'_{2,NC}\right) = -\frac{\gamma n (n-1)^2}{(1 + \gamma^2 n^3 + \gamma n (1+n))} < 0$$

which implies violation of condition (14).

Logarithmic Utility Case: Let $u(x) = \ln(x)$ and $D(P) = \frac{1}{2}\gamma P^2$. It is straightforward to show that the equilibrium strategies satisfy

$$E_{2,NC}(P_2) = \frac{\frac{\sqrt{4n+\gamma P_2^2}}{\sqrt{\gamma}} - P_2}{2n}$$

$$E_{2,Coop}(P_2) = \frac{\frac{\sqrt{4+\gamma P_2^2}}{\sqrt{\gamma}} - P_2}{2n}$$

Condition (14) becomes

$$G(P_1, n) \equiv \frac{-\sqrt{\gamma}P_1 + \sqrt{4n + \gamma P_1^2}}{-\sqrt{\gamma}P_1 + \sqrt{4 + \gamma P_1^2}} - \frac{(n+1) + \frac{(n-1)\sqrt{\gamma}P_1}{\sqrt{4n+\gamma P_1^2}}}{2} > 0 \tag{15}$$

In the limit case where P_1 tends to zero it is straightforward to see that we have

$$G(0,n) \equiv \sqrt{n} - \frac{n+1}{2} < 0 \text{ for } n > 1. \tag{16}$$

In Appendix B.4 we prove that (15) is violated in general, for $P_1 > 0$.

We therefore conclude that IEA type 4 is never environmentally friendly in the short run under LQ and log-utility specifications.

3.2 Long-run Impact

Since the impact of an anticipated IEA on first period emissions as well as second period emissions can be ambiguous, it is not clear that an IEA will necessarily result in an overall decrease of the stock of pollution in the long run, i.e. at the end of period 2. We prove below that all of the four types of IEAs share the feature that they are environmentally friendly in the long run, a feature which we think is indispensable given the *raison d'être* of IEAs. To this end we first establish a general result that applies to all potential IEAs. It provides two sufficiency conditions under which an IEA is environmentally friendly in the long run. For what is to follow, define $\mathcal{P}_2 \equiv \{P_2 \geq 0 | P_{3,C}(P_2) \equiv P_2 + nE_{2,C}(P_2) \geq P_{3,NC}\}$ for an initial stock P_1 where $P_{3,NC}$ is the equilibrium BAU pollution stock with initial pollution P_1. This is the set of period 2 pollution stock levels that leads to a higher equilibrium long-run pollution stock compared to BAU, when countries follow the given IEA.

Lemma 5. *Given an initial pollution stock $P_1 \geq 0$, let $P_{2,NC}$ and $P_{3,NC}$ denote the equilibrium BAU pollution levels. IEA $E_{2,C}(P_2)$ is environmentally friendly in the long run, i.e. $P_{3,C} \leq P_{3,NC}$, if the following two conditions are satisfied:*

1. $(1+(n-1)E'_{2,NC}(P_{2,NC}))D'(P_{3,NC}) \leq (1+(n-1)E'_{2,C}(P_2))D'(P_{3,C}) - E'_{2,C}(P_2)[u'(E_{2,C}(P_2)) - D'(P_{3,C})]$ *for all* $P_2 \in \mathcal{P}_2$, *where* $P_{3,C} \equiv P_2 + nE_{2,C}(P_2)$.

2. $P_2 + nE_{2,C}(P_2) \leq P_{3,NC}$ *for all* $P_2 \leq P_{2,NC}$.

Proof: See Appendix B.5.

Remark: It can be verified that under assumptions 1 and 2, if an IEA features $E_{2,C}(P_2) \leq E_{2,NC}(P_2)$ and $E'_{2,C}(P_2) \geq E'_{2,NC}(P_2)$ for all P_2, then it satisfies these two conditions. While the impact of the former condition is clear, the second inequality ensures that the given IEA features less "feedback effect" than under BAU.

The proof reveals that the first condition in lemma 5 is equivalent to the following statement: In any (potentially off-equilibrium) contingency in which IEA is *not* environmentally friendly in the long run, IEA must be environmentally friendly in the short run. The second condition states that whenever the IEA is environmentally friendly in the short run, it must be environmentally friendly in the long run.

Once these two conditions are satisfied by the given IEA, it is straightforward to show that it must be environmentally friendly in the long run.

At face value, all four IEAs are aimed at providing a cleaner environment relative to BAU. In general, whether an IEA actually is environmentally friendly, in the sense of Definition 1, depends critically on the countries' anticipatory response in period 1. It turns out all four IEAs reach their goals in the long run, independent of the anticipatory response in period 1.

We first establish the long-run impact of an IEA of Type 1 that sets an emission target $\bar{E}$ that satisfies $\bar{E} < E_{2,NC}(P_{2,NC})$.

Proposition 11. *Under assumption 1, IEA type 1 is environmentally friendly in the long run, i.e.* $P_{3,C} \leq P_{3,NC}$.

Proof: We refer to lemma 5 and show that the two conditions are satisfied. We have $E'_{2,C}(P_2) = 0$ for all P_2, therefore $(1 + (n-1)E'_{2,C}(P_2))D'(P_{3,C}) - E'_{2,C}(P_2)[u'(E_{2,C}(P_2)) - D'(P_{3,C})] = D'(P_{3,C})$ for all P_2.

Since $E'_{2,NC}(.) \leq 0$ by proposition 1, we have $(1 + (n-1)E'_{2,NC}(P_{2,NC}))D'(P_{3,NC}) \leq D'(P_{3,NC})$. Damage function is convex, and $P_2 \in \mathcal{P}_2$ implies $P_{3,C} \geq P_{3,NC}$, therefore $D'(P_{3,C}) \geq D'(P_{3,NC})$ for all $P_2 \in \mathcal{P}_2$. This proves that condition 1 in lemma 5 is met. Condition 2 is satisfied trivially since $E_{2,C}(P_2) \leq E_{2,NC}(P_{2,NC})$ for all P_2.■

Next, we turn to the long-run impact of IEA types 2, 3, and 4. The steps of the proof of the next proposition are very similar to those of the previous one, and the proof is therefore relegated to the appendix.

Proposition 12. *Under assumptions 1 and 2, IEA types 2, 3, and 4 are environmentally friendly in the long run for any initial pollution stock* P_1*, i.e.* $P_{3,C} \leq P_{3,NC}$.

Proof: See Appendix B.5.

Remark: Note that the conditions in lemma 5 are not necessary for an IEA to cause environmental degradation in the short run. In fact, each of the four IEAs we considered so far satisfies the properties in the lemma, yet, as shown in the previous sections, may not be environmentally friendly in the short run.

4 Welfare Analysis

We have shown that the impact of an anticipated IEA on first and second period emissions is ambiguous, therefore three questions related to welfare naturally arise: (i) "Do these IEAs lead to a welfare improvement over BAU?", (ii) if so, "For a given target of emissions in period 2, what type of IEA achieves the highest level of welfare?", and (iii) "Is there a clear welfare ranking between the four IEAs when their parameters are chosen optimally?". These questions are important for policy recommendations.

The first question posed is fundamental, and if the answer is negative, then the IEA should not even be considered. For IEA types 2 and 3, the answer is clearly positive because the planner can always choose $\varepsilon = \eta = 0$ to implement BAU level of emissions, hence the welfare level cannot decline if these parameters are chosen optimally. For IEA type 1, the answer is not obvious, because a fixed emissions target cannot implement the non-cooperative emissions profile in both periods in general. Nevertheless, we show below, that the highest welfare IEA type 1 can achieve is than welfare under BAU.

The second question is framed from the perspective of a policy maker whose main objective is to target a level of emissions for period 2. The answer to this question is not trivial because the same pollution levels can be reached using different IEAs by a suitable choice of parameters. We present some numerical results related to this question.

The third is related to constrained efficiency. In general, the planner will not be endowed with the policy tools to implement first-best emissions/welfare level. Rather, the planner solves the three problems of finding the welfare-maximizing IEA *within* the types $i \in \{1, 2, 3\}$. We address whether these three problems are well-defined and what qualitative/quantitative properties an optimal IEA of a given type possesses. We make some progress in this direction.

Let $(E^i_{1,C}, E^i_{2,C})$ and W^i_C be the equilibrium strategy and welfare level under the optimal IEA type $i \in \{1, 2, 3\}$ problem. IEA type 4 is not parametric, but for comparison, we will use the notation $(E^4_{1,C}, E^4_{2,C})$ and W^4_C to describe emissions and welfare levels under IEA type 4 respectively. Also define $P^i_{2,C} \equiv (1-\delta)(P_1 + nE^i_{1,C})$ for $i \in \{1, 2, 3, 4\}$ as the induced equilibrium second period pollution stock.

We first show that the optimal IEA type 1 problem is well-behaved and has a unique solution under our assumptions. Unfortunately, conditions even stronger than assumption 2 are necessary to warrant existence and uniqueness of solutions for the other types of IEAs. The difficulty arises from the fact that the choice set of the planner's problem may not be convex, a shortcoming encountered in many mechanism design problems. For theoretical results on IEA types 2 and 3, we will henceforth *assume* that there is a unique solution to the optimal IEA problem. Conditions ensuring existence and uniqueness are *always* satisfied under LQ setup, therefore, to ensure reliability, we will use numerical results based on this setup to compare the four types of IEAs.

The planner's problem for optimal IEA type 1 can be written as

$$W^1_C \equiv \max_{E_1, E_2 \geq 0} f(E_1, E_2) \tag{17}$$

subject to

$$E_1 \in \arg\max_{E \geq 0} \left\{ u(E) - D(P_1 + (n-1)E_1 + E) + \beta\big[u(E_2) - D\big((1-\delta)(P_1 + (n-1)E_1 + E) + nE_2\big)\big] \text{ given } E_1, E_2 \right\} \tag{18}$$

where function $f(.)$ is defined in equation (7). Since the objective in (18) is strictly concave in E, the first-order necessary condition is sufficient, therefore constraint (18) can be written as

$$u'(E_1) \leq D'(P_1 + nE_1) + \beta(1-\delta)D'\big((P_1 + nE_1)(1-\delta) + nE_2\big) \text{ with equality if } E_1 > 0. \tag{19}$$

Proposition 13. *Under assumptions 1 and 2,*

1. *There exists a unique solution to optimal IEA type 1 problem* (17).
2. *Welfare level under optimal IEA type 1, W_C^1, cannot be lower than the BAU level.*

Proof: See Appendix C.

Our second result compares welfare under IEA type 4, ex post first-best IEA, to welfare under optimal IEA type 1, 2 and 3 respectively. The following technical lemma applies to any parametric IEA, including IEAs type 1, 2 and 3 considered in this paper.

Lemma 6. *Suppose a parametric IEA $E_{2,C}(P_2;\theta)$ has the property that $E_{2,C}(P_{2,C}^4;\theta^*) = E_{2,C}^4$ for some θ^*. Then, (i) $E_{1,C}(P_1;\theta^*) = E_{1,C}^4$ and (ii) welfare under optimal IEA of this class satisfies $W_C \geq W_C^4$.*

Proof: See Appendix C.

Our claim follows as a corollary to this lemma since parameters $\bar{E}$, ε, η, for IEA type 1, 2, and 3 respectively, can always be chosen to satisfy the property $E_{2,C}(P_{2,C}^4;\theta^*) = E_{2,C}^4$ for some θ^*.

Corollary 2. *Welfare under IEA type 4 is dominated by welfare under optimal IEA types 1,2,3, i.e. $W_C^4 \leq W_C^i$, $i \in \{1,2,3\}$.*

In figure 1, we plot in the case of the LQ setup, welfare levels for IEA type 1, 2 and 3 as a function of *equilibrium* E_2: The parameters $\bar{E}$, ε, and η for IEAs type 1,2,3 are chosen such that equilibrium E_2 equals the corresponding value on the horizontal axis. For example, to generate a point on the graph of welfare of an IEA type 2, we determine, for a specific value of emissions in period 2 (say $\tilde{E}_2$), the value of the parameter ε that implements $E_{2,C}(P_2) = \tilde{E}_2$ in equilibrium. We then evaluate welfare for that particular value of ε. The same procedure is carried out for an IEA type 3. An IEA type 1 already has the emissions in period 2 as a parameter. We used the following parameter values to generate figure 1: $A = B = 1, \gamma = 1, \beta = 1, \delta = 0$ and $P_1 = 0$.[22]

The rightmost point in all plots represent BAU level $E_{2,NC}$. The welfare level achieved by IEA type 2 and type 3 at $E_2 = E_{2,NC}$ are obviously equal to BAU level since this level is induced by $\varepsilon = \eta = 0$.

[22] We conducted robustness checks for the insights we draw from figure 1 and that we discuss below. In the first numerical experiment we have used $A = B = 1, \gamma = 1, \delta = 0$ and $P_1 = 0$ and allowed the discount factor to take values in $\{0.85, 90, .95, 1\}$. In the second numerical experiment we set $A = B = 1, \gamma = 1, \beta = 1$ and $P_1 = 0$ and allowed the decay rate to take values in $\{0, 0.05, 0.10, 0.15\}$. In the third experiment we set $A = B = 1, \gamma = 1, \beta = 1$ and $\delta = 0$ and allowed P_1 to take values in $\{0, 0.10, 0.20, 0.30\}$.

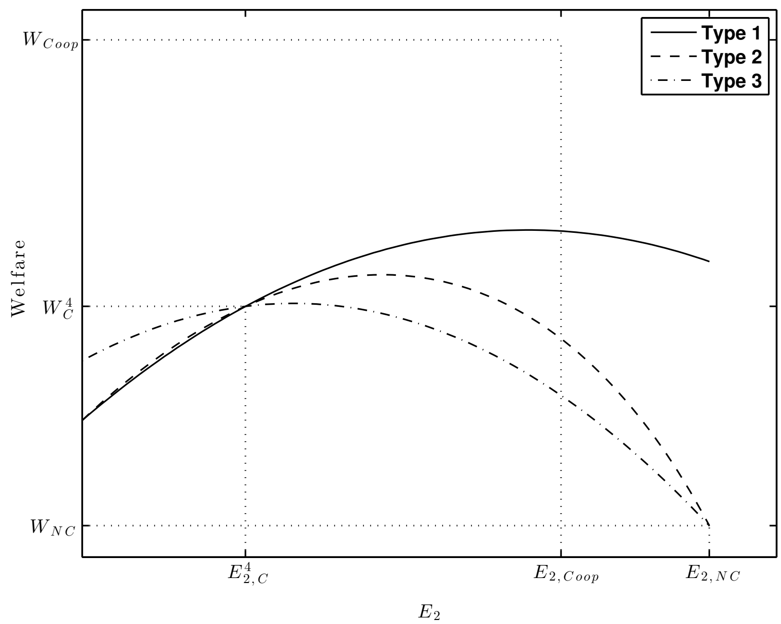

Figure 1: Welfare under different IEAs as a function of second period emissions.

Welfare under IEA type 1 at $\bar{E} = E_{2,NC}$ is higher than under BAU, because IEA type 1 effectively eliminates the feedback effect, which is partly responsible for inefficiently high emissions in period 1. For reference, we also indicate welfare and emissions levels under the first-best. Clearly, all four IEAs lead to a welfare improvement relative to BAU.

Another observation is that welfare functions cross at the same point, which represents emissions and welfare level under IEA type 4. In light of proof of lemma 6 and corollary 2, this is not a coincidence. IEA types 1,2,3 can all implement emissions (and welfare) under IEA type 4 by a suitable choice of parameter value.

A third observation is that W_C^1 is higher than W_C^2, W_C^3, and W_C^4, moreover, optimal emissions under types 1,2,3 are all higher than that under type 4. It turns out that this is always the case for environments in which the optimal IEA design problem is convex. In particular, it always holds in the LQ case. This follows from two results: (i) It can be shown that for IEA type 1, 2 and 3 implementing marginally higher emissions in period 2 than the $E_{2,C}^4$ level (by changing $\bar{E}, \varepsilon, \eta$ however necessary) improves welfare, and (ii) the following lemma, which states that if an optimal IEA features second period emissions higher than first-best levels, i.e. if $E_{2,C} \geq E_{2,Coop}(P_{2,C})$ holds, then switching to IEA type 1 always leads to a welfare improvement.[23]

[23] If the planner's problem is convex, the two claims imply IEA type 1 is unambiguously the best choice because the welfare functions are strictly concave with a single peak as in figure 1. A proof is available upon request from the authors for the LQ case.

Lemma 7. *Suppose the optimal choice of a parametric IEA,* $E_{2,C}(P_2;\theta^*)$*, features* $E'_{2,C}(P_2) \leq 0$ *for all* $P_2 \leq 0$*, and* $E_{2,C} \geq E_{2,Coop}(P_{2,C})$*. Then the induced welfare* W_C *is dominated by that achieved under the optimal IEA type 1, i.e.* $W_C^1 \geq W_C$*.*

Proof: See Appendix C.

Figure 1 also delivers insights into which type of IEA is better suited to achieve a given target of period 2 emissions. As noted above, given the three parametric IEAs, the same emissions cut can be implemented in alternative ways. Looking at the upper envelope of the welfare functions in figure 1 provides the planner with the best way in which a given emissions cut can be achieved, i.e. with the least welfare loss/highest welfare gain. For the case in figure 1, IEA type 1 is the best one to implement if the intended emissions cut is less than $\Delta \equiv E_{2,NC} - E^4_{2,C}$.[24] We would like to emphasize that an IEA type 1 is feasible only if commitment to a level of emissions in period 2 is possible. If players cannot commit to an emission level but can commit to a stock dependent emission policy for period 2, then the insight obtained from figure 1 is that an IEA type 2 is always preferable to an IEA type 3 if the intended emissions cut is less than Δ. To reduce emissions further, it is better to switch to an IEA type 3.

5 Optimal IEA: The General Case

We examine here the problem of designing an optimal cooperative agreement to announce, without restricting ourselves to the 4 types of agreements studied in the previous sections.

We say that an announcement $E_{2,C}(P_2)$ *implements emission levels* $(\bar{E}_1, \bar{E}_2)$ if (i) emission of $\bar{E}_1$ by all countries is a Nash equilibrium of the non-cooperative game played in period 1, given that they anticipate the given announcement, and (ii) $\bar{E}_2 = E_{2,C}((1-\delta)(P + n\bar{E}_1))$ holds. We note that after $E_{2,C}(P_2)$ is announced and period 1 unfolds, at the beginning of period 2, it would be welfare-improving to implement the emission that is optimal from that point of view: this corresponds to an announcement of an IEA type 4. Therefore the use of any announcement other than an IEA type 4 requires a commitment on behalf of the IEA members not to switch to the ex post first-best policy at the beginning of period 2. We assume that such a commitment is possible, similar to the standard assumption made in optimal fiscal policy literature in the Ramsey tradition. (See, for instance, Judd (1985) and Chamley (1986).) This is not too strong an assumption if the amount of emission reduction from the announced policy is smaller than the reduction of emission required under an IEA type 4. The rationale being that while it may be politically and logistically acceptable to implement a policy that is less stringent than originally anticipated, it is much more difficult for policy makers to switch to a more stringent policy, one that involves a bigger deviation from the BAU policy than initially announced.

[24] This can be proven in the LQ case. We conjecture that this remains true in the general case (provided assumptions 1 and 2 hold), however a proof is challenging due to potential lack of convexity of the planner's choice set.

We impose a mild restriction on the type of admissible IEAs for tractability. Let $\mathcal{C}$ be the space of continuously differentiable and monotone functions $g : \mathbb{R}_+ \to \mathbb{R}$ and let $\mathcal{F}$ be the space of functions $E_{2,C} : \mathbb{R}_+ \to \mathbb{R}_+$ that are of the form $E_{2,C}(P_2) = \max\{0, g(P_2)\}$ for some $g \in \mathcal{C}$. We will henceforth assume that our announcements satisfy $E_{2,C} \in \mathcal{F}$. Observe that these functions are continuously differentiable everywhere except for the point at which $g(P_2) = 0$. We argue that this is not very restrictive as a choice set, because first-best emissions strategies, BAU emissions strategies, and all IEA announcements we considered up to this point belong in $\mathcal{F}$.

The following proposition establishes two results. The first result provides necessary conditions for an announcement $E_{2,C} \in \mathcal{F}$ to implement a given emissions profile (E_1, E_2). The second result states that no announcement in $\mathcal{F}$ can implement the first-best level of emissions if both $(E_{1,Coop}, E_{2,Coop})$ are positive.

Proposition 14. *For given emissions $\bar{E}_1$, let $\bar{P}_2 = (1-\delta)(P_1 + n\bar{E}_1)$ denote the induced pollution stock at the beginning of period 2. Then,*

1. *If $E_{2,C} \in \mathcal{F}$ implements emission levels $(\bar{E}_1, \bar{E}_2)$ as a symmetric Nash equilibrium, then the following two conditions must hold:*

$$u'(\bar{E}_1) - D'(P_1 + n\bar{E}_1) + \beta(1-\delta)\left[E'_{2,C}(\bar{P}_2)(u'(\bar{E}_2) - nD'(\bar{P}_2 + n\bar{E}_2)) - D'(\bar{P}_2 + n\bar{E}_2)\right] \leq 0 \text{ with equality if } \bar{E}_1 > 0 \quad (20)$$

and

$$E_2 = E_{2,C}(\bar{P}_2)$$

2. *If $E_{1,Coop}, E_{2,Coop} > 0$, no announcement $E_{2,C} \in \mathcal{F}$ can implement first-best level of emissions and welfare.*

Proof:

1. The first inequality follows from the first-order necessary condition for optimality for a country i that takes profile $E_{1,-i} = \bar{E}_1$ of other countries as given, equation (10), explicitly accounting for the case of $\bar{E}_1 = 0$. The second equality follows from the definition of implementability.

2. Suppose $E_{2,C} \in \mathcal{F}$ implements first-best level of emissions $(\bar{E}_1, \bar{E}_2) = (E_{1,Coop}, E_{2,Coop}) > 0$. Then, $E_{2,C}$ satisfies (20) with equality at this point. Imposing second period optimality condition (9) (which holds with equality) on (20) leads to

$$u'(E_{1,Coop}) = D'(P_1 + nE_{1,Coop}) + \beta(1-\delta)D'((1-\delta)(P_1 + nE_{1,Coop}) + nE_{2,Coop})$$

This contradicts first period optimality condition in (9).■

Remark: To see why assumption $E_{1,Coop}, E_{2,Coop} > 0$ for part 2 is relevant, consider the case with $E_{1,Coop} = E_{2,Coop} = E_{1,NC} = E_{2,NC} = 0$. Obviously the trivial announcement $E_{2,C}(P_2) = 0$

implements first-best emissions levels, because in this particular case, the incentives of a planner and non-cooperative agents are perfectly aligned.

The following proposition establishes a more optimistic result which states that under mild conditions, *almost* any given emissions profile can be implemented using a linear announcement. For given emissions level $\bar{E}_1$, let $\bar{P}_2 = (1-\delta)(P_1 + n\bar{E}_1)$ denote the induced pollution stock at the beginning of period 2.

Proposition 15. *Suppose that u is such that $\lim_{E\downarrow 0} u'(E) = \infty$. For any $\bar{E}_1, \bar{E}_2 > 0$ such that $u'(\bar{E}_2) \neq nD'(\bar{P}_2 + n\bar{E}_2)$, there exists a unique pair (θ_0, θ_1) such that the announcement $E_{2,C}(P_2) = \max\{0, \theta_0 - \theta_1 P_2\}$ implements $(\bar{E}_1, \bar{E}_2)$.*

Proof: Given the emissions profile $\bar{E}_1, \bar{E}_2 > 0$, let

$$\theta_1 \equiv \frac{u'(\bar{E}_1) - D'(P_1 + n\bar{E}_1) - \beta(1-\delta)D'(\bar{P}_2 + n\bar{E}_2)}{\beta(1-\delta)(u'(\bar{E}_2) - nD'(\bar{P}_2 + n\bar{E}_2))} \tag{21}$$

$$\theta_0 \equiv \bar{E}_2 + \theta_1 \bar{P}_2$$

where $\bar{P}_2 = (1-\delta)(P_1 + n\bar{E}_1)$.

Under the given assumptions, these values are well-defined. We now claim that $E_{2,C}(P_2) = \max\{0, \theta_0 - \theta_1 P_2\}$ just constructed implements the given emissions levels. Given this announcement, and the assumption that $\lim_{E\downarrow 0} u'(E) = \infty$, first period emissions cannot induce a pollution stock P_2 such that $E_{2,C}(P_2) = 0$, hence we will ignore the "max" operator. Given the announcement $E_{2,C}(P_2) = \theta_0 - \theta_1 P_2$, and the emissions of all other countries $E_{1,-i} = \bar{E}_1$, the first-order necessary condition for period 1 emissions E_1 for a representative country i reads

$$u'(E_1) - D'(P_1 + (n-1)\bar{E}_1 + E_1) = \beta(1-\delta)\left[\theta_1(u'(E_{2,C}(P_2)) - nD'(P_2 + nE_{2,C}(P_2))) + D'(P_2 + nE_{2,C}(P_2))\right] \tag{22}$$

where $P_2 = (1-\delta)(P_1 + (n-1)\bar{E}_1 + E_1)$. Now observe that, given θ_1 in expression (21), $E_1 = \bar{E}_1$ is the *unique* solution to this equation. $E_1 = \bar{E}_1$ being a solution follows from the fact that (21) is a rearrangement of expression (22), and uniqueness follows from monotonicity of $u'(.)$ and $D'(.)$ by assumption 1. We conclude that the necessary conditions for implementability in proposition 14 is met. The second-order condition for this problem is also satisfied:

$$u''(\bar{E}_1) - D''(P_1 + n\bar{E}_1) + \beta(1-\delta)\left[\theta_1^2 u''(\bar{E}_2) - (1-\theta_1 n)^2 D''(\bar{P}_2 + n\bar{E}_2)\right] < 0.$$

Sufficiency for $\bar{E}_1$ being optimal is also met and this completes the proof.■

Remarks: (i) The inclusion of all potential emissions profiles in this proposition comes at the cost of the extra assumption that $\lim_{E\downarrow 0} u'(E) = \infty$. This assumption cannot be dispensed with, unless we restrict the set of admissible emissions profiles. For instance, profile $\bar{E}_1, \bar{E}_2 > 0$ could be so extreme that a country with quadratic utility would rather emit $E_2 = 0$. To do so, it might prefer deviating from $\bar{E}_1$ to change the pollution stock P_2 to the region where $E_{2,C}(P_2) = 0$. Clearly CRRA utility imposes no such restrictions. (ii) A linear function is not the unique announcement that implements any given $(\bar{E}_1, \bar{E}_2)$. Basically, any announcement that can target a given level and derivative/slope can implement the given levels, provided that necessary and sufficiency conditions for optimality are satisfied. The proof of this proposition is as simple as illustrating that there exists a unique solution to a system of two equations and two unknowns when the announcement is linear. Clearly, these properties are satisfied by a large class of non-linear announcements as well.

Proposition 15 has the striking implication that *almost* any emissions profile can be implemented by a linear announcement. Clearly, first-best solution is not in this domain, by proposition 14. In light of these findings, the natural question to ask is, "How well can we do in terms of welfare by choosing an announcement $E_{2,C} \in \mathcal{F}$?". It turns out the social planner can get *arbitrarily close* to fully cooperative emissions $(E_{1,Coop}, E_{2,Coop})$ and welfare levels. For this result, we omit the non-interesting cases where $E_{1,Coop} = E_{2,Coop} = 0$.[25] We would like to point out that the statement of the following proposition, and its proof is constructive, therefore it is has direct implications on the design and implementation of IEAs that are to be in effect in the future.

Proposition 16. *If $E_{1,Coop}(P_1), E_{2,Coop}(P_{2,Coop}) > 0$ holds, for any $\varepsilon > 0$ sufficiently small, there exist $\theta_0, \theta_1 > 0$ such that the announcement $E_{2,C}(P_2; \varepsilon) = \max\{0, \theta_0 - \theta_1 P_2\}$ implements the emission levels $(E_{1,Coop}, \frac{E_{2,Coop}}{1+\varepsilon})$ as a symmetric equilibrium.*

Proof: See Appendix D.

Remarks: (i) Note that assumption $\lim_{E \downarrow 0} u'(E) = \infty$ is not necessary for this result. (ii) Observe that Lemma 7 does not apply for the unconstrained case even when the subset of linear IEAs are considered: Lemma 7 *assumes* there exists an optimal announcement in the given class, it is violated here when the announcements are linear. Also, the second condition in Lemma 7 is violated.

Let $W(E_{2,C}; P_1)$ be the ex ante welfare level achieved using an announcement $E_{2,C}(P_2)$ when the initial pollution stock is P_1. Since welfare is continuous in emissions levels, the following result follows directly as a corollary of proposition 16. Even though no announcement can implement fully cooperative emissions (proposition 14), welfare levels arbitrarily close to first-best can be achieved.

Corollary 3. $\sup_{E_{2,C} \in \mathcal{F}} W(E_{2,C}; P_1) = W_{coop}(P_1)$ *for all* $P_1 \geq 0$.

Observe that the announcement described above is a very steep emission strategy. In fact, the proof reveals that as $\epsilon \to 0$, we have $\theta_0, \theta_1 \to \infty$. Second period emissions are potentially extremely high for an off-equilibrium contingency of $P_2 < \bar{P}_2 \equiv (1-\delta)(P + nE_{1,Coop})$. From a normative point of view, this is not very appealing. One way in which this problem can be mitigated is as

[25] See the remark after proposition 14.

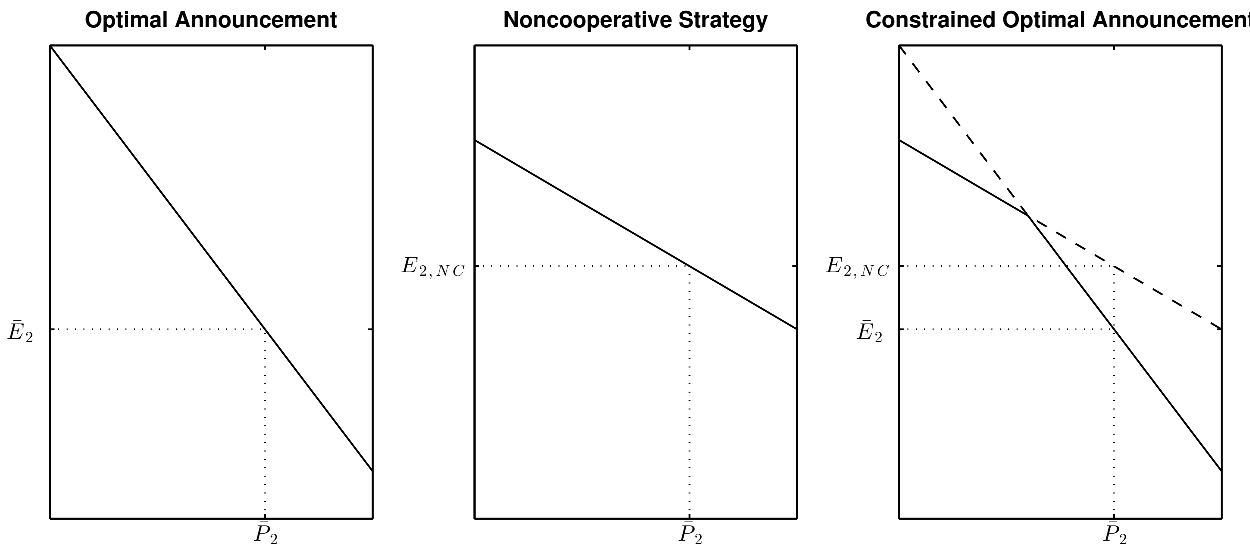

Figure 2: An example construction of $E_{2,C}(P_2) \in [0, E_{2,NC}(P_2)]$.

follows: Imagine that the planner faces an additional constraint that limits the emissions to be lower than non-cooperative emissions strategy and higher than zero, i.e. $E_{2,C}(P_2) \in [0, E_{2,NC}(P_2)]$ for all $P_2 \geq 0$. An example of this construction is shown in figure 2 with the announcement $E_{2,C}(P_2) \equiv \max\{0, \min\{E_{2,NC}(P_2), \theta_0 - \theta_1 P_2\}\}$. It is easy to show that when ε is sufficiently small, announcement $E_{2,C}(P_2)$ in Panel C of figure 2 still implements the emissions profile $\left(E_{1,Coop}, \frac{E_{2,Coop}}{1+\varepsilon}\right)$ as a symmetric equilibrium.

Lemma 8. *If $E_{1,Coop}(P), E_{2,Coop}(P_{2,Coop}) > 0$ holds, for $\varepsilon > 0$ sufficiently small, there exist $\theta_0, \theta_1 > 0$ such that the announcement $E_{2,C}(P_2) = \max\{0, \min\{E_{2,NC}(P_2), \theta_0 - \theta_1 P_2\}\}$ implements the emission levels $(E_{1,Coop}, \frac{E_{2,Coop}}{1+\varepsilon})$ as a symmetric equilibrium.*

Proof: See Appendix D.

6 Discussion and Concluding Remarks

The contribution of this paper is twofold. First, we have shown that strategic behavior can potentially exacerbate the tragedy of the commons in the pre-cooperation phase when there are delays in the agreement. Second, we have identified types of IEAs that can result in the attenuation of the tragedy of the commons problem in the pre-cooperation phase. In terms of policy recommendation, as far as the emission levels are concerned, an IEA that sets a target emission level or a percent cut in emission policy (in case the elasticity of marginal utility is smaller than unity) performs better than an IEA that implements a constant cut with respect to the BAU emission policy. We have compared within the LQ framework the welfare performance of the four different types of IEAs. Which IEA is preferable depends on the level emissions that one wishes to achieve in the second period. When period 2's emission target is above a certain threshold, an IEA type 1, that sets a target emission level, allows achieving a higher welfare than the other three types of IEAs examined. If players cannot commit to an emission level for period 2, when period 2's emission target is above a certain threshold, an IEA type 2 is preferable to an IEA type 3. Moreover the highest level of welfare achieved under an IEA type 1 is above the highest level of welfare achieved by any one of the other three IEAs. While we have established that in general no IEA can implement the first-best emissions, we proved existence of IEAs that can achieve a path of emissions (and therefore welfare) arbitrarily close to the first-best levels. Unfortunately the complexity of these IEAs increases substantially as we increase the number of periods that precedes the entry of the agreement into force, reducing the policy relevance of such schemes.

The incentive to emit more pollution in period 1, in anticipation of a future IEA can be reduced (or eliminated) by decreasing (or removing) the sensitivity of the emissions to the stock of pollution in period 2. Under an emission policy in period 2 that is downward sloping, each player has an extra benefit from emissions in period 1. Since an increase in private emissions results in an increase in the future stock of pollution, and in turn in a decrease of the other players' emissions. The strength of this feedback depends on how steep the emission policy is. IEAs of type 1 and 2 in affect not only reduce the level of emissions, but also the slope of the emission policy. In contrast, an IEA type 4, in the LQ and log-utility framework, implements a steeper emission policy. Not surprisingly, the feedback is exacerbated and each country responds by increasing its own current emissions.

In our framework, the damage incurred in period 2 is the result of the stock of pollution, after emissions in period 2 take place, $P_2 + nE_{2,NC}(P_2)$. One possible target that can be set by negotiators is to achieve a percentage cut in the *stock* of pollution, for instance, $(1-\varepsilon)(P_2 + nE_{2,NC}(P_2))$. Such an objective is equivalent to choosing an emission policy $E_{2,C}(P) = (1-\varepsilon) E_{2,NC}(P) - \frac{\varepsilon}{n}P$. The equivalent change in emission policy would imply a steeper environmental policy, a stronger feedback effect, and ultimately larger emissions in period 1 than under a percent cut in the BAU emission policy. Similarly, an IEA that aims to achieve a constant cut in the stock of pollution of the form $(P_2 + nE_{2,NC}(P_2)) - \eta$ is equivalent to adopting an emission policy $E_{2,C}(P) = E_{2,NC}(P) - \frac{\eta}{n}$, and would therefore have qualitatively the same effect as a constant cut with respect to the BAU emission policy, i.e., an increase in emissions in period 1 with respect to the BAU scenario. It is straightforward to establish this result as well as many of our results within an infinite time horizon model as we did for the case of proposition 6 where cooperation takes place from period 2 onwards.[26]

While we assumed throughout the paper, for simplicity, that the 'environmentally friendly' policy will be adopted in period 2 with certainty, our analysis extends to the case where there is uncertainty about whether the 'environmentally friendly' policy will actually be adopted (where the probability

[26] The steps are similar to the steps followed in the 2 period model. They are omitted, but are available from the authors upon request.

of adoption of the cooperative solution is positive and less than one). For example, in the U.S., a treaty is negotiated by the administration and ratified by the senate, where a sixty percent majority is required. In the U.K., the Constitutional Reform and Governance Act 2010 gave a statutory footing to the Ponsonby Rule, a convention that treaties be laid before Parliament before ratification.[27] The need of the approval of a legislative body often casts a significant amount of uncertainty on the actual ratification of a treaty signed by the executive branch of government.[28] Our analysis shows that even a treaty that may end up being rejected at the ratification stage will have an impact on the current policy (i.e., policy during the phase between adoption and the date of the ratification decision) as long as there is a positive probability of ratification in the future. Depending on the nature of the externality, and on the form of the anticipated future 'friendly' policy, the expectation of the ratification alone can alleviate or worsen the tragedy of the commons.

A natural extension of our analysis is to move our stylized model closer to reality by introducing country heterogeneity. In the case of climate change for example, while some countries are mildly affected by global warming, others will experience catastrophic damages. The reaction of each group of countries to an announced agreement is likely to differ, and reciprocally countries heterogeneity may modify the type of IEA that countries may announce.

We believe that the lessons learned from this analysis extend to dynamic stock public good games in general (see e.g., Fershtman and Nitzan (1991)). Anticipating a future agreement that will increase the provision of a public good with a respect to a business-as-usual policy may result in an increase or a decrease of current contributions of the public good. The optimistic message of this paper is that it is possible to have agreements over contribution policies for the future than can alleviate the tragedy of the commons in the present. To this end, a careful examination of the slopes of the contribution policy and the BAU policy is required.

[27] In Australia, the Federal Government does not need parliamentary approval to enter a binding treaty. However, legislation by Federal parliament is needed for the implementation of treaties.

[28] For example: the Kyoto Protocol was signed but not ratified by the U.S.; the initial 1973 Convention for the Prevention of Pollution From Ships (Marpol 73) did not come into force for lack of ratifications.

References

Barrett, S., (1998), "On the Theory and Diplomacy of Environmental Treaty-Making," *Environmental and Resource Economics,* Vol. 11 (3-4): 317-333.

Barrett, S., (2005). The Theory of International Environmental Agreements. In K.-G. Mäler and J. Vincent (eds.), Handbook of Environmental Economics, Vol. 3, Amsterdam: Elsevier, pp. 1457-1516.

Beccherle, J. and J. Tirole (2011), "Regional initiatives and the cost of delaying binding climate change agreements," *Journal of Public Economics*, 95: 1339-1348.

Benchekroun, H. and A. Ray Chaudhuri, (2014) "Transboundary pollution and clean technologies," forthcoming, *Resource and Energy Economics*.

Benchekroun, H., (2003), "Unilateral production restrictions in a dynamic duopoly," *Journal of Economic Theory*, 111(2): 214-239.

Calvo, E. and S. J. Rubio (2012), "Dynamic Models of International Environmental Agreements: A Differential Game Approach," *International Review of Environmental and Resource Economics*, 6: 289-339.

Chamley, Christophe (1986), "Optimal Taxation of Capital Income in General Equilibrium with Infinite Lives." *Econometrica* 54(3): 607-22.

Congleton, R. D. (2001), "Governing the global commons: the political economy of international environmental treaties and institutions." In Schulze, G.G. and Ursprung, H.W. (eds.) Globalization and the environment. pp 241-263. New York: Oxford University Press.

Dockner, E. and A. A. Haug, (1990), "Tariffs and Quotas under Dynamic Duopolistic Competition," *Journal of International Economics*, 29: 147-59.

Dockner, E. and A. A. Haug, (1991), "Closed-Loop Motive for Voluntary Export Restraints," *Canadian Journal of Economics*, 24: 679-85.

Dockner, E. J., S. Jorgensen, N.V. Long, and G. Sorger (2000). "Differential games in economics and management science," Cambridge University Press.

Eichner, T. and R., Pethig (2011), "Carbon Leakage, The Green Paradox, And Perfect Future Markets," *International Economic Review*, vol. 52(3): 767-805.

Fershtman, C. and S. Nitzan (1991), "Dynamic Voluntary Provision of Public Goods," *European Economic Review*, 35: 1057-67.

Friend, I. and M. E. Blume (1975), "The Demand for Risky Assets," *The American Economic Review*, 65(5): 900-22.

Germain, M., P. Toint, H. Tulkens, and A. de Zeeuw (2003), "Transfers to sustain dynamic core-theoretic cooperation in international stock pollutant control," *Journal of Economic Dynamics and Control*, 28: 79-99.

Grafton, Q., Kompas, T., and N. V. Long (2012), "Substitution between biofuels and fossil fuels: is there a Green Paradox?," *Journal of Environmental Economics and Management*, Vol. 64(3): 328-341.

Hall, R. E. (1988), "Intertemporal Substitution in Consumption," *Journal of Political Economy*, 96(2): 339-357

Harstad, B. (2012), "Climate Contracts: A Game of Emissions, Investments,Negotiations, and Renegotiations," *Review of Economic Studies*, 79: 1527-1557.

Harstad, B. (2014), "The Market for Conservation and Other Hostages," working paper, University of Oslo.

Harstad, B. (2015), "The Dynamics of Climate Agreements," forthcoming in the *Journal of European Economics Association*.

Havranek, T., R. Horvath, Z. Irsova, and M. Rusnak (2015), "Cross-Country Heterogeneity in Intertemporal Substitution," *Journal of International Economics*, 96: 100-118.

Hoel, M. (1991),"Global Environmental Problems: The Effects of Unilateral Actions taken by one country," *Journal of Environmental Economics and Management,* 20: 55-70.

Hoel, M. (1992), "International environmental conventions: the case of uniform reductions of emissions," *Environmental and Resource Economics*, 2: 141-59.

Holt, C. A. and S. K. Laury, "Risk Aversion and Incentive Effects," *The American Economic Review* 92(5): 1644-1655.

Jørgensen, S., G. Martin-Herran and G. Zaccour (2010) "Dynamic Games in the Economics and Management of Pollution," *Environmental Modeling and Assessment*, 15: 433-467.

Judd, Kenneth L. (1985), "Redistributive taxation in a simple perfect foresight model." *Journal of Public Economics* 28(1): 59-83.

Long, N. V., (2011), "Dynamic games in the economics of natural resources: A survey," *Dynamic Games and Applications,* 1(1): 115-148.

Long, N. V., (2014), "The Green Paradox in Open Economies," CESifo Working Paper No. 4639.

Oates, W. E. and P. R. Portney (2003), "The political economy of environmental policy." In Mäler, K. G. and Vincent, J. R., (eds.). Handbook of Environmental Economics. pp. 325–354. Amsterdam: Elsevier.

Petrosyan, L., and G. Zaccour (2003), "Time-consistent Shapley value allocation of pollution cost reduction," *Journal of Economic Dynamics and Control,* 27: 381-398.

Van der Ploeg, F. and C. Withagen (2012), "Is there really a green paradox?," *Journal of Environmental Economics and Management*, 64: 342-363.

Van der Ploeg, F., and C. Withagen (2013), "Global warming and the green paradox," OxCarre, Discussion Paper 116.

Sinn, H. W. (2008), "Public Policies against Global Warming: A Supply-Side Approach," *International Tax and Public Finance*, 15(4): 360-394.

Sinn, H. W. (2012), *The Green Paradox*, MIT Press.

Smulders, S., Y. Tsur, and A. Zemel (2012), "Announcing climate policy: Can a green paradox arise without scarcity?," *Journal of Environmental Economics and Management*, 64: 64-376.

Van der Werf, E. and C. Di Maria (2012), "Imperfect environmental policy and polluting emissions:

the Green Paradox and beyond," *International Review of Environmental and Resource Economics*, 6 (2): 153-194.

A Proofs for Section 2

Proposition 1. *Under assumption 1, there exists a unique non-cooperative Nash equilibrium for period 2. The equilibrium is symmetric, and the strategy* $E_{2,NC} : \mathbb{R}_+ \to \mathbb{R}_+$ *solves*

$$u'\left(E_{2,NC}(P_2)\right) \leq D'\left(P_2 + nE_{2,NC}(P_2)\right) \text{ with equality if } E_{2,NC}(P_2) > 0. \tag{23}$$

In addition, $E_{2,NC}$ *has the following properties:*

1. *It is continuously differentiable and* $E'_{2,NC}(P_2) \in [-\frac{1}{n}, 0]$ *for all* $P_2 > 0$*, except at* $\bar{P}_2 \equiv \lim_{E\downarrow 0}(D')^{-1}\big(u'(E)\big)$ *if this point is well-defined.*

2. *If, in addition, assumption 2 holds, strategy* $E_{2,NC}(P_2)$ *is convex, and the value of following this strategy,* $V_2(P_2) \equiv u(E_{2,NC}(P_2)) - D(P_2 + nE_{2,NC})$*, is decreasing and concave in* P_2.

Proof: Under assumption 1, objective in problem (1), $f(E)$, is strictly concave in E and $-f(.)$ is coercive in the convex constraint set $E \geq 0$. Hence, there exists a unique $E_{2,i}$ that solves this problem and the following Kuhn-Tucker first-order necessary condition is sufficient.

$$u'\left(E_{2,i}\right) \leq D'\left(P_2 + E_{2,i} + \sum_{k\in I\setminus\{i\}} E_{2,k}\right) \text{ with equality if } E_{2,i} > 0.$$

Then strategy $E_{2,NC}(P_2)$ is a symmetric Nash equilibrium if and only if condition (23) is satisfied. First, we will show that if there exists such a function, it must be unique. Then we will establish existence using a constructive argument. Obviously, there cannot be a Nash equilibrium that is not symmetric, because the marginal damage component is identical for all countries, and both marginal utility and damage functions are strictly monotone.

Uniqueness: Suppose the solution exists, but is not unique, and that there are two solutions that differ at some P_2: $E \equiv E_{2,NC}(P_2) > \tilde{E}_{2,NC}(P_2) \equiv \tilde{E}$. Since $\tilde{E} \geq 0$, $E > 0$ must hold. Then

$$D'(P_2 + nE) = u'(E) \leq u'(\tilde{E}) \leq D'(P_2 + n\tilde{E})$$

where the first inequality follows from concavity of $u(.)$. Since $D(.)$ is convex, we have $E \leq \tilde{E}$, a contradiction.

Existence: Let $g(x; P_2) \equiv u'(x) - D'(P_2 + nx)$. This function is strictly decreasing over $\mathbb{R}_{++}$ by assumption 1 and it is easy to verify that $\lim_{x\to\infty} g(x; P_2) < 0$ by assumption 1 part 3. Define $\bar{P}_2 \equiv \lim_{E\downarrow 0}(D')^{-1}\big(u'(E)\big)$. This value is well-defined if $u'(0) > 0$, otherwise $\bar{P}_2 = \infty$. For all $P_2 < \bar{P}_2$, $\lim_{x\downarrow 0} g(x; P_2) \geq 0$, hence, there exists a unique $x^*(P_2) > 0$ that solves $g(x^*(P_2); P_2) = 0$ by intermediate value theorem. Now let

$$E_{2,NC}(P_2) = \begin{cases} 0 & \text{for } P_2 \geq \bar{P}_2 \\ x^*(P_2) & \text{otherwise .} \end{cases}$$

This function is well-defined, and by construction, it satisfies the symmetric Nash equilibrium condition. This proves existence.

1. Continuity and differentiability over the region in which $E_{2,NC}(P_2) > 0$ follows from continuity and differentiability of its closed-form (explicit) inverse $P_2(E_{2,NC}) = (D')^{-1}(u'(E_{2,NC})) - nE_{2,NC}$. It is also easy to check that $\lim_{P_2 \uparrow \bar{P}_2} E_{2,NC}(P_2) = 0$, hence there is no jump discontinuity at $P_2 = \bar{P}_2$.

   At an interior solution, i.e. when $P_2 < \bar{P}_2$, we have $u''(E_{2,NC})E'_{2,NC} = D''(.)(1 + nE'_{2,NC})$. By assumption 1 item 2, we have $E'_{2,NC} \leq 0$, since otherwise left-hand side and right-hand side would have different signs. Also observe that $E'_{2,NC} \geq -\frac{1}{n}$ must hold, otherwise right-hand side would be negative and left-hand side would be positive. For the region in which $E_{2,NC} = 0$, $E'_{2,NC} = 0$ must hold, so the claim is trivially satisfied.

2. *Strategy is convex:* At an interior solution $u''(E_{2,NC})E'_{2,NC} = D''(.)(1 + nE'_{2,NC})$ holds. Assume that $u'(.)$ is convex and $D'(.)$ is concave. Suppose, to get a contradiction, that $E'_{2,NC}$ is decreasing in an open neighborhood of some P_2. Since $E'_{2,NC} \leq 0$, left-hand side would unambiguously increase as P_2 goes up infinitesimally. Looking at the right-hand side, since $E'_{2,NC}(P_2) \in [-\frac{1}{n}, 0]$ holds, the term $(1 + nE'_{2,NC})$ and $D''(P_2 + nE_{2,NC})$ both unambiguously decrease as P_2 goes up infinitesimally. Then, the equality cannot be preserved as P_2 is varied. Convexity of the strategy function is robust to corner solutions, because $\max\{0, f(x)\}$ is a convex function whenever $f(x)$ is a convex function.

   Value function is decreasing and concave: Differentiate the value function and impose the first-order necessary condition to get

   $$V'_2(P_2) = -\big(1 + (n-1))E'_{2,NC}(P_2)\big)D'(P_2 + nE_{2,NC}(P_2))$$

   Clearly this derivative exists whenever $E'_{2,NC}(P_2)$ exists. By assumption 1 and $E'_{2,NC}(P_2) \in [-\frac{1}{n}, 0]$, $V'_2(P_2) \leq 0$. Differentiating the value function a second time wherever a derivative exists, we obtain

   $$V''_2(P_2) = -(n-1)E''_{2,NC}(P_2)D'(P_2+nE_{2,NC}(P_2))-[1+(n-1))E'_{2,NC}(P_2)]\big(1+nE'_{2,NC}(P_2)\big)D''(P_2+nE_{2,NC}(P_2)).$$ [29]

   Since the strategy is convex, $E''_{2,NC}(P_2) \geq 0$, and since $E'_{2,NC}(P_2) \in [-\frac{1}{n}, 0]$, we have $[1 + (n-1))E'_{2,NC}(P_2)]\big(1 + nE'_{2,NC}(P_2)\big) \geq 0$. Then $V''_2(P_2) \leq 0$ as we wanted to show.■

Proposition 4. *Under assumption 1,*

1. *Long run equilibrium pollution stock satisfies $P_{3,Coop} \leq P_{3,NC}$ for all $P_1 \geq 0$.*
2. *Short run equilibrium pollution stock satisfies $P_{2,Coop} \leq P_{2,NC}$ for all $P_1 \geq 0$.*

Proof:

[29] Value function being twice continuously differentiable is not necessary to establish this result.

1. Suppose, this is not the case and $P_{3,Coop} > P_{3,NC} \geq 0$ holds. First, we show that this assumption implies $E_{1,Coop} > E_{1,NC}$ must hold. If $E_{2,Coop} = 0$, this condition holds trivially since otherwise $P_{3,Coop} > P_{3,NC}$ is violated. Consider the non-trivial case of $E_{2,Coop} > 0$: Using propositions 1 and 3, we have

   $$u'(E_{2,Coop}) = nD'(P_{3,Coop}) \geq nD'(P_{3,NC}) \geq nu'(E_{2,NC})$$

   where the first inequality follows from convexity of $D(.)$ and our assumption on pollution stocks and the second inequality follows from equation (2). Since $u(.)$ is concave and $n > 1$, $E_{2,Coop} \leq E_{2,NC}$ must hold. But then inequality $E_{1,Coop} > E_{1,NC} \geq 0$ is satisfied, otherwise $P_{3,Coop} > P_{3,NC}$ is violated.

   Next, we show that $E_{1,Coop} > E_{1,NC}$ and $P_{3,Coop} > P_{3,NC}$ jointly lead to a contradiction. Using the first-order necessary condition in proposition 3, we have

   $$\begin{aligned} u'(E_{1,Coop}) &= n\Big(D'\big(P_1 + nE_{1,Coop}\big) + \beta(1-\delta)D'(P_{3,Coop})\Big) \\ &\geq n\Big(D'\big(P_1 + nE_{1,NC}\big) + \beta(1-\delta)D'(P_{3,NC})\Big) \end{aligned} \tag{24}$$

   where the inequality follows from convexity of $D(.)$, $E_{1,Coop} > E_{1,NC} \geq 0$ and $P_{3,Coop} > P_{3,NC}$. Using proposition 2, we have

   $$\begin{aligned} u'\left(E_{1,NC}\right) &\leq D'\left(P_1 + nE_{1,NC}\right) + \beta\left(1-\delta\right)\left(1 + \left(n-1\right)E'_{2,NC}\right)D'\left(P_{3,NC}\right) \\ &\leq D'\left(P_1 + nE_{1,NC}\right) + \beta\left(1-\delta\right)D'\left(P_{3,NC}\right) \end{aligned} \tag{25}$$

   where the inequality follows from $E'_{2,NC} \leq 0$. Combining inequalities (24) and (25), we have $u'(E_{1,Coop}) \geq nu'(E_{1,NC})$. Since $u(.)$ is concave and $n > 1$, $E_{1,Coop} \leq E_{1,NC}$ must hold, a contradiction.

2. Suppose, to get a contradiction, that $P_{2,Coop} > P_{2,NC}$, or equivalently, $E_{1,Coop} > E_{1,NC} \geq 0$. Since $u(.)$ is concave, $u'(E_{1,Coop}) \leq u'(E_{1,NC})$ must hold. Then using propositions 2 and 3, we obtain

   $$\begin{aligned} u'(E_{1,Coop}) = n\Big(D'\big(P_1 + nE_{1,Coop}\big) + \beta(1-\delta)D'(P_{3,Coop})\Big) &\leq u'(E_{1,NC}) \\ &\leq D'\left(P_1 + nE_{1,NC}\right) + \beta\left(1-\delta\right)\left(1 + \left(n-1\right)E'_{2,NC}\right)D'\left(P_{3,NC}\right) \\ &\leq D'\left(P_1 + nE_{1,NC}\right) + \beta\left(1-\delta\right)D'\left(P_{3,NC}\right) \end{aligned}$$

   where the last inequality follows from $E'_{2,NC} \leq 0$. Collecting terms, we get

   $$\begin{aligned} nD'\big(P_1 + nE_{1,Coop}\big) - D'\left(P_1 + nE_{1,NC}\right) &\leq \beta(1-\delta)\big(D'\left(P_{3,NC}\right) - nD'(P_{3,Coop})\big) \\ (n-1)D'\left(P_1 + nE_{1,NC}\right) &\leq \beta(1-\delta)\big(u'(E_{2,NC}) - u'(E_{2,Coop})\big) \end{aligned}$$

   where the inequality follows from convexity of $D(.)$, $E_{1,Coop} > E_{1,NC}$, the first-order necessary conditions for period 2 under BAU, ($u'(E_{2,NC}) = D'(P_{3,NC})$, equal since $E_{2,NC} > E_{2,Coop} \geq 0$, otherwise $P_{3,Coop} \leq P_{3,NC}$ in part 1 is violated) and first-best ($u'(E_{2,Coop}) \leq nD'(P_{3,Coop})$). Since left-hand side is non-negative, $u'(E_{2,NC}) \geq u'(E_{2,Coop})$ holds, which implies $E_{2,Coop} \geq E_{2,NC}$. Along with $E_{1,Coop} > E_{1,NC}$, we must have $P_{3,Coop} > P_{3,NC}$, this contradicts part 1 of this proposition. ■

B Proofs for Section 3

B.1 IEA Type 1

Lemma 1. *For an IEA of type 1, a decrease in the threshold $\bar{E}$ unambiguously results in an increase in current emissions and therefore an increase in $P_{2,C}$.*

Proof: Total differentiation with respect to $\bar{E}$ gives

$$u''(E_{1,C})\frac{dE_{1,C}}{d\bar{E}} - D''(P_1+nE_{1,C})\,n\frac{dE_{1,C}}{d\bar{E}} - \beta(1-\delta)D''\left[(P_1+nE_{1,C})(1-\delta)+n\bar{E}\right]\left(n(1-\delta)\frac{dE_{1,C}}{d\bar{E}}+n\right)=0$$

or

$$\begin{aligned}&\frac{dE_{1,C}}{d\bar{E}}\left(u''(E_{1,C}) - D''(P_1+nE_{1,C})\,n\frac{dE_{1,C}}{d\bar{E}} - \beta(1-\delta)D''[(P_1+nE_{1,C})(1-\delta)+n\bar{E}]n(1-\delta)\right)\\ =\;&\beta(1-\delta)D''\left[(P_1+nE_{1,C})(1-\delta)+n\bar{E}\right]n\end{aligned}$$

that is

$$\frac{dE_{1,C}}{d\bar{E}} = \frac{\beta(1-\delta)D''\left[(P_1+nE_{1,C})(1-\delta)+n\bar{E}\right]n}{u''(E_{1,C}) - D''(P_1+nE_{1,C})\,n\frac{dE_{1,C}}{d\bar{E}} - \beta(1-\delta)D''\left[(P_1+nE_{1,C})(1-\delta)+n\bar{E}\right]n(1-\delta)} < 0$$

Therefore a decrease in $\bar{E}$ unambiguously results in an increase of $E_{1,C}$.■

Lemma 2. *An IEA of type 1 that features $\bar{E} = E_{2,NC}(P_{2,NC})$ is environmentally friendly in the short run.*

Proof: When $\bar{E} = E_{2,NC}(P_{2,NC})$, we have

$$u'(E_{1,C}) - D'(P_1+nE_{1,C}) - \beta(1-\delta)D'\left[(P_1+nE_{1,C})(1-\delta)+nE_{2,NC}(P_{2,NC})\right]=0$$

with

$$\left[u'(E_{2,NC}(P_2)) - nD'(P_2+nE_{2,NC}(P_2))\right]E'_{2,NC}(P_2) > 0$$

whereas in the non-cooperative scenario

$$\begin{aligned}u'(E_{1,NC}) - D'(P_1+nE_{1,NC}) =&-\beta(1-\delta)\left\{u'(E_{2,NC}(P_2)) - nD'(P_2+nE_{2,NC}(P_2))E'_{2,NC}(P_2)\right\}\\ &+\beta(1-\delta)D'(P_2+nE_{2,NC}(P_2))\\ =&0\end{aligned}$$

Let

$$h(E) = u'(E) - D'(P_1 + nE) - \beta(1-\delta) D'\left((P_1 + nE)(1-\delta) + n\bar{E}\right)$$

we have $h' < 0$ and

$$h(E_{1,C}) = 0$$

and at $\bar{E} = E_{2,NC}(P_{2,NC})$ we have

$$h(E_{1,NC}) = -\beta(1-\delta)\{u'(E_{2,NC}(P_2)) - nD'(P_2 + nE_{2,NC}(P_2))\} E'_{2,NC}(P_2) < 0$$

and therefore

$$E_{1,C} < E_{1,NC}. \blacksquare$$

Setting $\bar{\mathbf{E}}$ close to zero leads to an increase in emissions:

The non-cooperative equilibrium emission $E_{1,NC}$ is solution to

$$\begin{aligned} u'(E_{1,NC}) - D'(P_1 + nE_{1,NC}) &= \beta(1-\delta) D'(P_{2,NC} + nE_{2,NC}(P_{2,NC})) \\ &-\beta(1-\delta)\{u'(E_{2,NC}(P_{2,NC})) - nD'(P_{2,NC} + nE_{2,NC}(P_{2,NC}))\} E'_{2,NC}(P_{2,NC}) \end{aligned} \tag{26}$$

where $P_{2,NC} = (P_1 + nE_{1,NC})(1-\delta)$. We know that in period 2 we have

$$u'(E_{2,NC}(P_2)) - D'(P_2 + nE_{2,NC}(P_2)) = 0$$

so

$$u'(E_{1,NC}) - D'(P_1 + nE_{1,NC}) = \beta(1-\delta) D'(P_{2,NC} + nE_{2,NC}(P_{2,NC}))\left[1 + (n-1) E'_{2,NC}(P_{2,NC})\right] \tag{27}$$

Under an IEA that sets a target emissions level, the equilibrium emission level in period 1, $E_{1,C}$, is solution to

$$u'(E_{1,C}) - D'(P_1 + nE_{1,C}) = \beta(1-\delta) D'\left(P_{2,C} + n\bar{E}\right) \tag{28}$$

where $P_{2,C} = (P_1 + nE_{1,C})(1-\delta)$.

First we note that that for $\beta = 0$ we have $E_{1,C} = E_{1,NC}$ and $P_{2,C} = P_{2,NC}$. Moreover total differentiation of (27) with respect β gives

$$\begin{aligned} & [u''(E_{1,NC}) - D''(P_1 + nE_{1,NC})] n \frac{dE_{1,NC}}{d\beta} \\ = \; & (1-\delta) D'(P_{2,NC} + nE_{2,NC}(P_{2,NC}))\left[1 + (n-1) E'_{2,NC}(P_{2,NC})\right] \\ & + \beta(1-\delta) \frac{d\left\{D'(P_{2,NC} + nE_{2,NC}(P_{2,NC}))\left[1 + (n-1) E'_{2,NC}(P_{2,NC})\right]\right\}}{dP_2} \frac{dP_{2NC}}{dE_{1,NC}} \frac{dE_{1,NC}}{d\beta} \end{aligned}$$

which, when $\beta = 0$, simplifies into

$$[u''(E_{1,NC}) - D''(P_1 + nE_{1,NC})] n \left.\frac{dE_{1,NC}}{d\beta}\right|_{\beta=0} = (1-\delta) D'(P_{2,NC} + nE_{2,NC}(P_{2,NC}))\left[1 + (n-1) E'_{2,NC}(P_{2,NC})\right].$$

Similarly, total differentiation of (28) with respect β gives

$$\left[u''\left(E_{1,C}\right)-D''\left(P_1+nE_{1,C}\right)\right]n\frac{dE_{1,C}}{d\beta}=\left(1-\delta\right)D'\left(P_{2,C}+n\bar{E}\right)+\beta\left(1-\delta\right)\frac{dD'\left(P_{2,C}+\bar{E}\right)}{dE_{1,NC}}\frac{dE_{1,C}}{d\beta}$$

which, when $\beta=0$, yields

$$\left[u''\left(E_{1,C}\right)-D''\left(P_1+nE_{1,C}\right)\right]n\left.\frac{dE_{1,C}}{d\beta}\right|_{\beta=0}=\left(1-\delta\right)D'\left(P_{2,C}+n\bar{E}\right)$$

So

$$\begin{aligned}&\left[u''\left(E_{1,NC}\right)-D''\left(P_1+nE_{1,NC}\right)\right]n\left(\left.\frac{dE_{1,C}}{d\beta}\right|_{\beta=0}-\left.\frac{dE_{1,NC}}{d\beta}\right|_{\beta=0}\right)\\=\;&\left(1-\delta\right)\left\{D'\left(P_{2,C}+n\bar{E}\right)-D'\left(P_{2,NC}+nE_{2,NC}\left(P_{2,NC}\right)\right)\left[1+\left(n-1\right)E'_{2,NC}\left(P_{2,NC}\right)\right]\right\}\end{aligned}$$

or

$$\begin{aligned}&\left[u''\left(E_{1,NC}\right)-D''\left(P_1+nE_{1,NC}\right)\right]n\left(\left.\frac{dE_{1,C}}{d\beta}\right|_{\beta=0}-\left.\frac{dE_{1,NC}}{d\beta}\right|_{\beta=0}\right)\\=\;&\left(1-\delta\right)\left\{D'\left(P_{2,C}+n\bar{E}\right)-D'\left(P_{2,NC}+nE_{2,NC}\left(P_{2,NC}\right)\right)\left[1+\left(n-1\right)E'_{2,NC}\left(P_{2,NC}\right)\right]\right\}\end{aligned}$$

for $\bar{E}=0$, since $u''-D''<0$, we have

$$\begin{aligned}&sign\left(\lim_{\delta\uparrow 1}\left(\left.\frac{dE_{1,C}}{d\beta}\right|_{\beta=0}-\left.\frac{dE_{1,NC}}{d\beta}\right|_{\beta=0}\right)\right)\\=\;&-sign\left(\lim_{\delta\uparrow 1}\left\{D'\left(P_{2,C}\right)-D'\left(P_{2,NC}+nE_{2,NC}\left(P_{2,NC}\right)\right)\left[1+\left(n-1\right)E'_{2,NC}\left(P_{2,NC}\right)\right]\right\}\right)\end{aligned}$$

We know that

$$1+nE'_{2,NC}\left(P_2\right)>0$$

and therefore, for $E'_{2,NC}\left(P_2\right)<0$, we have

$$1+\left(n-1\right)E'_{2,NC}\left(P_2\right)>0.$$

Moreover in the limit case where $\delta\uparrow 1$ we have full depreciation of the stock of pollution and therefore $P_2\to 0$. We have

$$\begin{aligned}&\lim_{\delta\uparrow 1}\left\{D'\left(P_{2,C}\right)-D'\left(P_{2,NC}+nE_{2,NC}\left(P_{2,NC}\right)\right)\left[1+\left(n-1\right)E'_{2,NC}\left(P_{2,NC}\right)\right]\right\}\\=\;&\lim_{\delta\uparrow 1}\left\{D'\left(0\right)-D'\left(nE_{2,NC}\left(0\right)\right)\left[1+\left(n-1\right)E'_{2,NC}\left(0\right)\right]\right\}\end{aligned}$$

Since $E_{2,NC}\left(0\right)>0$ and $\left[1+\left(n-1\right)E'_{2,NC}\left(0\right)\right]>0$, when $D'\left(0\right)=0$ we have

$$\lim_{\delta\uparrow 1}\left\{D'\left(P_{2,C}\right)-D'\left(P_{2,NC}+nE_{2,NC}\left(P_{2,NC}\right)\right)\left[1+\left(n-1\right)E'_{2,NC}\left(P_{2,NC}\right)\right]\right\}<0$$

and thus

$$\lim_{\delta\uparrow 1}\left(\left.\frac{dE_{1,C}}{d\beta}\right|_{\beta=0}-\left.\frac{dE_{1,NC}}{d\beta}\right|_{\beta=0}\right)>0.$$

Recall that $E_{1,C}=E_{1,NC}$ when $\beta=0$. We can therefore conclude that, when $D'\left(0\right)=0$, for $\bar{E}=0$, there exists $\bar{\beta}>0,\bar{\delta}\in(0,1)$ such that

$$E_{1,C}-E_{1,NC}>0\text{ for }0<\beta<\bar{\beta}\text{ and }1>\delta>\bar{\delta}.\blacksquare$$

B.2 IEA Type 2

Proposition 6. *A type 2 IEA that features a marginal proportional reduction in emissions is environmentally friendly in the short run if and only if* $\sigma \equiv -\frac{E_{2,NC}(P_2)u''(E_{2,NC}(P_2))}{u'(E_{2,NC}(P_2))} < 1$

Proof:

We have

$$\begin{aligned}\frac{R'(\varepsilon)}{-\beta(1-\delta)} &= -u'\left\{(1-\varepsilon)E_{2,NC}(P_2)\right\}E'_{2,NC}(P_2) - E_{2,NC}(P_2)u''\left\{(1-\varepsilon)E_{2,NC}(P_2)\right\}(1-\varepsilon)E'_{2,NC}(P_2)\\ &\quad -\left(-nE'_{2,NC}(P_2)\right)D'\left(P_2+n(1-\varepsilon)E_{2,NC}(P_2)\right)\\ &\quad -\left[1+n(1-\varepsilon)E'_{2,NC}(P_2)\right]\left(-nE_{2,NC}(P_2)\right)D''\left(P_2+n(1-\varepsilon)E_{2,NC}(P_2)\right)\end{aligned}$$

The impact of a marginal change in ε in the neighborhood of $\varepsilon = 0$ is given by

$$\begin{aligned}\frac{R'(0)}{\beta(1-\delta)} &= \left[u'(E_{2,NC}(P_2)) - nD'(P_2+nE_{2,NC}(P_2))\right]E'_{2,NC}(P_2)\\ &\quad +E_{2,NC}(P_2)u''(E_{2,NC}(P_2))E'_{2,NC}(P_2)\\ &\quad -n\left(1+nE'_{2,NC}(P_2)\right)E_{2,NC}(P_2)D''(P_2+nE_{2,NC}(P_2))\end{aligned}$$

Using expression (3) gives

$$\begin{aligned}\frac{R'(0)}{\beta(1-\delta)} &= \left[u'(E_{2,NC}(P_2)) - nD'(P_2+nE_{2,NC}(P_2))\right]E'_{2,NC}(P_2)\\ &\quad +E_{2,NC}(P_2)\left[u''(E_{2,NC}(P_2))E'_{2,NC}(P_2) - nu''(E_{2,NC}(P_2))E'_{2,NC}(P_2)\right]\end{aligned}$$

or

$$\begin{aligned}\frac{R'(0)}{\beta(1-\delta)} &= \left[u'(E_{2,NC}(P_2)) - nD'(P_2+nE_{2,NC}(P_2))\right]E'_{2,NC}(P_2)\\ &\quad +E_{2,NC}(P_2)(1-n)u''(E_{2,NC}(P_2))E'_{2,NC}(P_2)\end{aligned}$$

that is

$$\frac{R'(0)}{\beta(1-\delta)E'_{2,NC}(P_2)} = \underbrace{u'(E_{2,NC}(P_2)) - nD'(P_2+nE_{2,NC}(P_2))}_{<0} + \underbrace{E_{2,NC}(P_2)(1-n)u''(E_{2,NC}(P_2))}_{>0} \tag{29}$$

Note that $u'(E_{2,NC}(P_2)) - nD'(P_2+nE_{2,NC}(P_2)) < 0$ from $E_{2,NC}(P_2)$ larger than first-best emissions which solve $u'(E) - nD'(P_2+nE) = 0$.

The sign of $R'(0)$ is therefore undetermined.

The condition (29) may be rewritten as

$$\frac{R'(0)}{\beta(1-\delta)} = -(n-1)E'_{2,NC}(P_2)u'(E_{2,NC}(P_2))(1-\sigma)$$

where

$$\sigma \equiv -\frac{E_{2,NC}\left(P_2\right)u''\left(E_{2,NC}\left(P_2\right)\right)}{u'\left(E_{2,NC}\left(P_2\right)\right)}$$

The sign of $R'(0)$ depends on the elasticity of marginal utility. It is positive if and only if $\sigma < 1$.■

Infinite-Horizon Extension

Under infinite horizon, BAU value function satisfies the following recursive relation under the optimal policy $E_{NC}(P)$

$$V(P) = u(E_{NC}(P)) - D(P + nE_{NC}(P)) + \beta V\big((1-\delta)(P + nE_{NC}(P))\big)$$

The first-order necessary condition (ignoring a corner solution) for each country is

$$u'(E_{NC}(P)) - D'(P + nE_{NC}(P)) = -\beta(1-\delta)V'(P') \tag{30}$$

where P' denotes next period pollution stock and

$$V'(P') = u'(E_{NC}(P'))E'_{NC}(P') - (D'(P + nE_{NC}(P')) - \beta(1-\delta)V'(P''))(1 + nE'_{NC}(P'))$$

Differentiating expression (30) equality with respect to P, we obtain an expression for the derivative of the strategy:

$$u'(.)E'_{NC}(P) = (1 + nE'_{NC}(P))D''(.) - \beta(1-\delta)V''(P') \tag{31}$$

Now consider the announcement $E_C(P) = (1-\varepsilon)E_{NC}(P)$ to be effective from next period onwards. Just as we did in the two-period case, we can write the following analogous conditions

$$u'(E_{NC}(P)) - D'(P + nE_{NC}(P)) = -\beta(1-\delta)V'(P'; E_C) \equiv R(\varepsilon)$$

where

$$R(\varepsilon; P') \equiv -\beta(1-\delta)\left\{u'[(1-\varepsilon)E_{NC}(P')](1-\varepsilon)E'_{NC} - [1+n(1-\varepsilon)E'_{NC}][D'(P'+(1-\varepsilon)E_{NC}(P')) - \beta(1-\delta)V'(P'')]\right\}$$

and $V'(P'; E_C)$ represents the derivative of the value function in which all occurances of $E_{NC}(P)$ are replaced with $(1-\varepsilon)E_{NC}(P')$. Obviously, as in the two-period case, the sign of $R'(0)$ determines whether a marginal increase in ε leads to an increase or decrease in emissions today, relative to BAU. Differentiating $R(\varepsilon)$ with respect to ε, imposing $\varepsilon = 0$, the first-order condition (30), and the derivative condition (31), we obtain the simplified version

$$\frac{R'(0)}{\beta(1-\delta)} = -(n-1)E'_{NC}u'(.)\left(\frac{u''(.)E_{NC}}{u'(.)} + 1\right) = -(n-1)E'_{NC}u'(.)(1-\sigma).$$

Therefore the same conclusion holds under infinite horizon.

Lemma 4. *There exists a unique $\bar{\gamma} > 0$ such that $\Omega < 0$ for $\gamma > \bar{\gamma}$ and $\Omega > 0$ for $0 < \gamma < \bar{\gamma}$.*

Proof: We first establish existence of $\bar{\gamma}$ such that $\Omega = 0$. This follows from the fact that $\Omega(\gamma = 0) = 1 > 0$, $\lim_{\gamma\to\infty}\Omega = -\infty$ and that Ω is continuous.

We now establish uniqueness of $\bar{\gamma}$. The function Ω can have at most three roots. From $\Omega(\gamma = 0) = 1 > 0$ we can infer that Ω cannot have two positive roots only. Suppose Ω has three positive roots then one of the following two statements must hold Ω is strictly decreasing strictly convex in the neighborhood of $\gamma = 0$

$$\frac{d\Omega}{d\gamma} = \left(-1 + \beta(1-\delta)^2 + 2\delta\right)n - 3n^3\gamma^2 + 2n\gamma\left(\beta(1-\delta)^2 + n(2\delta-3)\right)$$

$$\frac{d^2\Omega}{d\gamma^2} = -6n^3\gamma + 2n\left(\beta(1-\delta)^2 + n(2\delta-3)\right)$$

At $\gamma = 0$, $\left.\frac{d\Omega}{d\gamma}\right|_{\gamma=0} = \left(\beta(1-\delta)^2 + 2\delta - 1\right)n < 0$

$$\left.\frac{d^2\Omega}{d\gamma^2}\right|_{\gamma=0} = 2n\left(\beta(1-\delta)^2 + n(2\delta-3)\right) > 0$$

$$\left.\frac{d^2\Omega}{d\gamma^2}\right|_{\gamma=0} = 2n\left(\frac{1}{n}\left.\frac{d\Omega}{d\gamma}\right|_{\gamma=0} + (n-1)2\delta - 3n + 1\right)$$

The term $(n-1)2\delta - 3n + 1$ is a strictly increasing function of δ for $n > 1$ therefore, since $\delta \in [0,1]$ is no larger than $(n-1)2 - 3n + 1 = -n - 1 < 0$. Therefore if $\left.\frac{d\Omega}{d\gamma}\right|_{\gamma=0} \leq 0$ we must have $\left.\frac{d^2\Omega}{d\gamma^2}\right|_{\gamma=0} < 0$, i.e. Ω is strictly concave in the neighborhood of $\gamma = 0$. This rules out the existence of three positive roots of Ω. This completes the proof.■

B.3 IEA Type 3

Proposition 9. *An IEA of type 3 that features a marginal cut in emissions by all players is not environmentally friendly in the short run.*

Proof: We have

$$\frac{Q'(\eta)}{-\beta(1-\delta)} = -u''\left(E_{2,NC}(P_2) - \eta\right)E'_{2,NC}(P_2) + n\left(1 + nE'_{2,NC}(P_2)\right)D''\left(P_2 + nE_{2,NC}(P_2) - n\eta\right)$$

and thus

$$\frac{Q'(0)}{-\beta(1-\delta)} = -u''\left(E_{2,NC}(P_2)\right)E'_{2,NC}(P_2) + n\left(1 + nE'_{2,NC}(P_2)\right)D''\left(P_2 + nE_{2,NC}(P_2)\right)$$

From (3)

$$u''(E_{2,NC})E'_{2,NC} - \left(1 + nE'_{2,NC}\right)D''(P_2 + nE_{2,NC}) = 0 \tag{32}$$

so

$$\frac{Q'(0)}{-\beta(1-\delta)} = (n-1)u''\left(E_{2,NC}(P_2)\right)E'_{2,NC}(P_2) > 0$$

or

$$Q'(0) < 0$$

and therefore in the neighborhood of $\eta = 0$ we have

$$Q(\eta) < Q(0) \text{ for } \eta > 0$$

and thus

$$E_{1,C}(P_1) > E_{1,NC}(P_1) \text{ for all } P_1.\blacksquare$$

B.4 IEA Type 4

Proposition 10. *If IEA type 4 is environmentally friendly in the short run, then the following condition must hold:*

$$\frac{u'(E_{2,Coop}(P_{2,Coop}))}{u'(E_{2,NC}(P_{2,NC}))} - n\left(1 + (n-1)E'_{2,NC}(P_{2,NC})\right) > 0.$$

Proof: Using equation (9) we have

$$u'(E_{1,C}(P_1)) - D'(P_1 + nE_{1,C}(P_1)) - \beta(1-\delta)\frac{u'(E_{2,Coop}(P_2))}{n} = 0$$

Recall that for the non-cooperative equilibrium we have

$$u'(E_{1,NC}(P_1)) - D'(P_1 + nE_{1,NC}(P_1)) - \beta(1-\delta)\left(1 + (n-1)E'_{2,NC}\right)u'(E_{2,NC}(P_2)) = 0$$

If $E_{1,NC}(P_1) > E_{1,C}(P_1)$ then, since $u'' - D'' < 0$,

$$u'(E_{1,NC}(P_1)) - D'(P_1 + nE_{1,NC}(P_1)) < u'(E_{1,C}(P_1)) - D'(P_1 + nE_{1,C}(P_1))$$

that is

$$\beta(1-\delta)\left(1 + (n-1)E'_{2,NC}\right)u'(E_{2,NC}(P_2)) < \beta(1-\delta)\frac{u'(E_{2,Coop}(P_2))}{n}$$

or

$$\frac{u'(E_{2,Coop}(P_{2,Coop}))}{u'(E_{2,NC}(P_{2,NC}))} - n\left(1 + (n-1)E'_{2,NC}(P_{2,NC})\right) > 0. \tag{33}$$

where

$$P_{2,NC} = P_1 + nE_{1,NC}(P_1)$$

and

$$P_{2,Coop} = P_1 + nE_{1,C}(P_1)$$

The above condition (33) is a necessary condition for $E_{1,NC}(P_1) > E_{1,C}(P_1)$.$\blacksquare$

Lemma 5. *For $P_1 \geq 0$, condition (14) is never satisfied, i.e., $G(P_1, n) < 0$*

Proof: For clarity of notation, we will let $P \equiv P_1$. Without loss of generality, let $\gamma = 1$. We have

$$G(P,n) = \frac{-P+\sqrt{4n+P^2}}{-P+\sqrt{4+P^2}} - \frac{n+1+\frac{(n-1)P}{\sqrt{4n+P^2}}}{2} \tag{34}$$

we rewrite

$$\frac{-P+\sqrt{4n+P^2}}{-P+\sqrt{4+P^2}} = 1 + \frac{-\sqrt{4+P^2}+\sqrt{4n+P^2}}{-P+\sqrt{4+P^2}}$$

or

$$\frac{-P+\sqrt{4n+P^2}}{-P+\sqrt{4+P^2}} = 1 + \frac{4n-4}{\left(-P+\sqrt{4+P^2}\right)\left(\sqrt{4+P^2}+\sqrt{4n+P^2}\right)}$$

We can therefore write that

$$\begin{aligned} &\left(-P+\sqrt{4+P^2}\right)\left(\sqrt{4+P^2}+\sqrt{4n+P^2}\right) \\ = \quad & 4+P^2-P\sqrt{4+P^2}+\sqrt{4+P^2}\sqrt{4n+P^2} > 4+2\sqrt{4n+P^2} \end{aligned}$$

because we have $P^2 - P\sqrt{4+P^2} < 0$ and $\sqrt{4+P^2} > 2$.

We thus have

$$\frac{-P+\sqrt{4n+P^2}}{-P+\sqrt{4+P^2}} < 1 + \frac{4n-4}{4+2\sqrt{4n+P^2}}$$

which implies that

$$G(P,n) < 1 + \frac{4n-4}{4+2\sqrt{4n+P^2}} - \frac{n+1+\frac{(n-1)P}{\sqrt{4n+P^2}}}{2} \tag{35}$$

or

$$G(P,n) < 1 - \frac{n+1}{2} + \frac{4n-4}{4+2\sqrt{4n+P^2}} - \frac{(n-1)P}{2\sqrt{4n+P^2}} \tag{36}$$

We can rewrite this inequality as

$$G(P,n) < (n-1)\left(-\frac{1}{2} - \frac{P}{2\sqrt{4n+P^2}} + \frac{4}{4+2\sqrt{4n+P^2}}\right) \tag{37}$$

or

$$G(P,n) < -\frac{(n-1)}{2\sqrt{P^2+4n}\,(P^2+4n-4)}\left\{\left(P^2+4+4n\right)\sqrt{P^2+4n} - 16n - 4P + 4Pn - 4P^2 + P^3\right\}. \tag{38}$$

where $A \equiv \left(P^2+4+4n\right)\sqrt{P^2+4n} - 16n - 4P + 4Pn - 4P^2 + P^3$.

We now show that the expression A is positive. Indeed, we have

$$A > \bar{A} \equiv -\left(P^2+4+4n\right)\sqrt{P^2+4n} - 16n - 4P + 4Pn - 4P^2 + P^3$$

moreover the product $A\bar{A}$ gives

$$A\bar{A} = -8P^5 - 4P^4n - 64P^3n + 32P^3 - 32P^2n^2 + 32P^2n - 128Pn^2 + 128Pn - 64n^3 + 128n^2 - 64n$$

or

$$A\bar{A} = -8P^5 - 4P^4n - 32P^3(2n-1) - 32P^2n(n-1) - 128Pn(n-1) - 64n(n-1)^2 < 0$$

Since $A\bar{A} < 0$ and $A > \bar{A}$ this implies that $A > 0$ and therefore from equation (38), we have that $G(P,n) < 0$.■

B.5 Long-Run Impact

Lemma 5. *Given an initial pollution stock $P_1 \geq 0$, let $P_{2,NC}$ and $P_{3,NC}$ denote the equilibrium BAU pollution levels. IEA $E_{2,C}(P_2)$ is environmentally friendly in the long run, i.e. $P_{3,C} \leq P_{3,NC}$, if the following two conditions are satisfied:*

1. $(1+(n-1)E'_{2,NC}(P_{2,NC}))D'(P_{3,NC}) \leq (1+(n-1)E'_{2,C}(P_2))D'(P_{3,C}) - E'_{2,C}(P_2)[u'(E_{2,C}(P_2)) - D'(P_{3,C})]$ *for all* $P_2 \in \mathcal{P}_2$, *where* $P_{3,C} \equiv P_2 + nE_{2,C}(P_2)$.

2. $P_2 + nE_{2,C}(P_2) \leq P_{3,NC}$ *for all* $P_2 \leq P_{2,NC}$.

Proof: Fix $P_1 \geq 0$, and assume, to get a contradiction, that the two conditions above hold, but that $P_{3,C} > P_{3,NC}$. Obviously, $P_{2,C} \in \mathcal{P}_2$, therefore condition 1 implies

$$\begin{aligned}[1+(n-1)E'_{2,NC}(P_{2,NC})]D'(P_{3,NC}) \leq\ &[1+(n-1)E'_{2,C}(P_{2,C})]D'(P_{3,C}) \\ &- E'_{2,C}(P_{2,C})[u'(E_{2,C}(P_{2,C})) - D'(P_{3,C})]\end{aligned}$$

We have $E_{1,C} > 0$, otherwise $P_{2,C} = 0$ and condition 2 contradicts $P_{3,C} > P_{3,NC}$. The first-order condition (11), the first-order condition for BAU in proposition 2, and the inequality above jointly imply

$$u'(E_{1,C}) - D'(P_1 + nE_{1,C}) \geq u'(E_{1,NC}) - D'(P_1 + nE_{1,NC})$$

or equivalently,

$$u'(E_{1,C}) - u'(E_{1,NC}) \geq D'(P_1 + nE_{1,C}) - D'(P_1 + nE_{1,NC})$$

Utility is concave, and damage function is convex, therefore $E_{1,C} \leq E_{1,NC}$, and $P_{2,C} \leq P_{2,NC}$ hold. Then condition 2 implies $P_{3,C} \leq P_{3,NC}$, a contradiction.■

Proposition 12. *Under assumptions 1 and 2, IEA types 2, 3, and 4 are environmentally friendly in the long run for any initial pollution stock P_1, i.e. $P_{3,C} \leq P_{3,NC}$.*

Proof: We verify the two conditions in lemma 5 are satisfied for any $P_1 \geq 0$. Fix some $P_1 \geq 0$ and take any $P_2 \in \mathcal{P}_2$.

IEA type 2: Consider the IEA $E_{2,C}(P_2) = (1-\varepsilon) E_{2,NC}(P_2)$ where $\varepsilon \in [0,1]$. By straightforward algebra, we can write

$$\begin{aligned}&(1+(n-1)E'_{2,C}(P_2))D'(P_{3,C}) - E'_{2,C}(P_2)[u'(E_{2,C}(P_2)) - D'(P_{3,C})] \\ &\qquad = (1+(n-1)E'_{2,NC}(P_2))D'(P_{3,C}) - (n-1)\epsilon E'_{2,NC}(P_2)D'(P_{3,C}) \\ &\qquad\quad - (1-\epsilon)E'_{2,NC}(P_2)(u'(E_{2,C}(P_2)) - D'(P_{3,C})) \\ &\qquad \geq (1+(n-1)E'_{2,NC}(P_2))D'(P_{3,C}) \\ &\qquad \geq (1+(n-1)E'_{2,NC}(P_{2,NC}))D'(P_{3,NC})\end{aligned}$$

where the first inequality holds due to the fact that last two terms are negative for all $P_2 \geq 0$: If $E_{2,NC}(P_2) > 0$, this follows from $E'_{2,NC} \leq 0$ and $u'(E_{2,C}(P_2)) \geq u'(E_{2,NC}(P_2)) = D'(P_2 + nE_{2,NC}(P_2)) \geq D'(P_2 + nE_{2,C}(P_2)) \equiv D'(P_{3,C})$. Otherwise $E'_{2,NC}(P_2) = 0$ holds so the claim is trivially true. The last inequality follows from convexity of $E_{2,NC}$ established in proposition 1, convexity of damage function, and $P_{3,C} \geq P_{3,NC}$ due to $P_2 \in \mathcal{P}_2$. We conclude that condition 1 in lemma 5 is satisfied.

To show that condition 2 is satisfied, take any $P_2 \leq P_{2,NC}$. We have $P_2 + n(1-\epsilon)E_{2,NC}(P_2) \leq P_2 + nE_{2,NC}(P_2) \leq P_{2,NC} + nE_{2,NC}(P_{2,NC}) = P_{3,NC}$ where the last inequality follows from the fact that final pollution stock is increasing in P_2 under BAU. (See the remark after proposition 1).

IEA type 3: Consider the case where countries anticipate an IEA in period 2 that features a constant cut in emissions $\eta > 0$. By straightforward algebra, we can write

$$\begin{aligned}
&(1+(n-1)E'_{2,C}(P_2))D'(P_{3,C}) - E'_{2,C}(P_2)(u'(E_{2,C}(P_2)) - D'(P_{3,C})) \\
=&(1+(n-1)E'_{2,NC}(P_2))D'(P_{3,C}) - E'_{2,C}(P_2)(u'(E_{2,C}(P_2)) - D'(P_{3,C})) \\
\geq&(1+(n-1)E'_{2,NC}(P_{2,NC}))D'(P_{3,NC}) - E'_{2,C}(P_2)(u'(E_{2,C}(P_2)) - D'(P_{3,C})) \\
\geq&(1+(n-1)E'_{2,NC}(P_{2,NC}))D'(P_{3,NC})
\end{aligned}$$

the first inequality follows from convexity of $E_{2,NC}$ (proposition 1), $P_{3,C} \geq P_{3,NC}$ (due to $P_2 \in \mathcal{P}_2$) and convexity of $D(.)$. The last inequality follows from $u'(E_{2,C}(P_2)) - D'(P_{3,C}) \geq 0$ for all P_2, this is the case because $E_{2,C}(P_2) \leq E_{2,NC}(P_2)$ for all P_2. Verification of the claim that the second condition in lemma 5 follows the same steps in the proof of Type 2.

IEA type 4: Consider the case where countries anticipate that there will be full cooperation in period 2, i.e. $E_{2,C}(P_2) = E_{2,Coop}(P_2)$. By straightforward algebra, we can write

$$\begin{aligned}
&(1+(n-1)E'_{2,C}(P_2))D'(P_{3,C}) - E'_{2,C}(P_2)(u'(E_{2,C}(P_2)) - D'(P_{3,C})) \\
=&D'(P_{3,C}) - E'_{2,C}(P_2)(u'(E_{2,C}(P_2)) - nD'(P_{3,C})) \\
=&D'(P_{3,C}) \\
\geq&(1+(n-1)E'_{2,NC}(P_{2,NC}))D'(P_{3,NC})
\end{aligned}$$

the first equality follows by collecting terms. The second equality follows from the first-order necessary condition for optimality under full cooperation, $u'(E_{2,C}(P_2)) = nD'(P_{3,C})$. The inequality follows from $E'_{2,NC} \leq 0$, convexity of $D(.)$ and $P_2 \in \mathcal{P}_2$. Hence condition 1 is satisfied. We leave the verification of condition 2 to readers as the steps are virtually identical to those used for IEA type 2.■

C Proofs for Section 4

Proposition 13. *Under assumptions 1 and 2,*

1. *There exists a unique solution to optimal IEA type 1 problem* (17).

2. *Welfare level under optimal IEA type 1, W_C^1, cannot be lower than the BAU level.*

Proof: Consider the following relaxed problem:

$$\tilde{W}_C^1 \equiv \max_{E_1,E_2 \geq 0} f(E_1, E_2) \tag{39}$$

subject to

$$u'(E_1) \leq D'(P_1 + nE_1) + \beta(1-\delta)D'\big((P_1 + nE_1)(1-\delta) + nE_2\big) \tag{40}$$

Clearly constraint set (40) includes (19) and therefore $\tilde{W}_C^1 \geq W_C^1$.

1. Under assumption 1, the objective in problem (39), $f(E_1, E_2)$ is jointly strictly concave in (E_1, E_2) and $-f(.)$ is coercive, as in proposition 3. The constraint set (40) is closed, non-empty, and under the assumptions $u'''(.) \geq 0$ and $D'''(.) \leq 0$, it is convex.[30] Therefore, there exists a solution $(\tilde{E}_1, \tilde{E}_2)$, it is unique, and moreover, Kuhn-Tucker necessary conditions of the associated Lagrangian are sufficient.

   Next, we show that $(\tilde{E}_1, \tilde{E}_2)$ belongs to the constraint set (19). This shows that relaxed problem achieves its maximum inside the original constraint and hence it must be a solution to problem (17). There are two possibilities for $(\tilde{E}_1, \tilde{E}_2)$:

   (i) The constraint (40) does not bind: Then it must be that $(\tilde{E}_1, \tilde{E}_2) = (E_{1,Coop}, E_{2,Coop})$ since latter is the unique solution to the unconstrained problem as shown in proposition 3. It is easy to show that $E_{1,Coop} = \tilde{E}_1 = 0$. If this is not the case,

$$\begin{aligned} u'(\tilde{E}_1) &= n\Big(D'(P_1 + n\tilde{E}_1) + \beta(1-\delta)D'\big[(P_1 + n\tilde{E}_1)(1-\delta) + n\tilde{E}_2\big]\Big) \\ &> \Big(D'(P_1 + n\tilde{E}_1) + \beta(1-\delta)D'\big[(P_1 + n\tilde{E}_1)(1-\delta) + n\tilde{E}_2\big]\Big) \end{aligned}$$

   The first equality holds since this is the first-order necessary condition for the fully cooperative solution. The inequality follows from $n > 1$. But then $(\tilde{E}_1, \tilde{E}_2)$ violates constraint (40). This proves $\tilde{E}_1 = 0$ and that $(\tilde{E}_1, \tilde{E}_2)$ satisfies (19).

   (ii) The constraint (40) binds: This case is trivial, obviously $(\tilde{E}_1, \tilde{E}_2)$ satisfies (19).

   We have thus shown that problem (17) has a unique solution and it can be characterized by solving the relaxed problem (39).

2. Fully non-cooperative/BAU emissions $(E_{1,NC}, E_{2,NC})$ solve

$$u'(E_{1,NC}) - \Big(D'(P_1 + nE_{1,NC}) + \beta(1-\delta)[1 + (n-1)E'_{2,NC}(P_2)]D'\big[(P_1 + nE_{1,NC})(1-\delta) + nE_{2,NC}\big]\Big) \leq 0$$

   with equality when $E_{1,NC} > 0$. Since $E'_{2,NC}(P_2) \leq 0$, this implies

$$u'(E_{1,NC}) \leq D'(P_1 + nE_{1,NC}) + \beta(1-\delta)D'\big[(P_1 + nE_{1,NC})(1-\delta) + nE_{2,NC}\big]$$

[30] It is non-empty, because it contains BAU emissions $(E_{1,NC}, E_{2,NC})$ as shown in part 2 of the proof.

But then $(E_{1,NC}, E_{2,NC})$ belongs to constraint set (40), so the welfare level achieved under optimal IEA type 1, W_C^1 cannot be below BAU level.■

Lemma 6. *Suppose a parametric IEA $E_{2,C}(P_2;\theta)$ has the property that $E_{2,C}(P^4_{2,C};\theta^*) = E^4_{2,C}$ for some θ^*. Then, (i) $E_{1,C}(P_1;\theta^*) = E^4_{1,C}$ and (ii) welfare under optimal IEA of this class satisfies $W_C \geq W_C^4$.*

Proof: It suffices to show that under parameter $\theta = \theta^*$, this IEA implements $(E^4_{1,C}, E^4_{2,C})$ and therefore obtains a welfare level of W_C^4. This implies welfare cannot deteriorate if θ is chosen optimally. Consider the first-order necessary condition for $E_{1,C}$ when $E_{2,C}(P_2;\theta^*)$ is announced, and impose symmetry:

$$u'(E_{1,C}) - D'(P_1 + nE_{1,C}) + \beta(1-\delta)\big[E'_{2,C}(.)(u'(E_{2,C}(.)) - nD'(P_2 + nE_{2,C}(.))) - D'(P_2 + nE_{2,C}(.))\big] \leq 0$$

with equality if $E_{1,C} > 0$. We will assume that this problem is convex so that it has a unique solution. It is easy to see that $(E^4_{1,C}, E^4_{2,C})$ satisfies this condition: We first guess that $E_{1,C} = E^4_{1,C}$, and $P_2 = P^4_{2,C}$ and then verify it. By assumption, $E_{2,C}(P^4_{2,C};\theta^*) = E^4_{2,C}$. Since $E^4_{2,C}$ is the equilibrium under ex post first-best, we have $u'(E_{2,C}(.)) - nD'(P^4_{2,C} + nE_{2,C}(.)) = 0$. Then, the condition above simplifies to

$$u'(E_{1,C}) - D'(P_1 + nE_{1,C}) - \beta(1-\delta)D'(P^4_{2,C} + nE^4_{2,C})\big] \leq 0$$

with equality if $E_{1,C} > 0$. Observe that $E_{1,C} = E^4_{1,C}$ satisfies this condition since this is the first-order necessary condition for $E_{1,C}$ under IEA type 4.■

Lemma 7. *Suppose the optimal choice of a parametric IEA, $E_{2,C}(P_2;\theta^*)$, features $E'_{2,C}(P_2) \leq 0$ for all $P_2 \geq 0$, and $E_{2,C} \geq E_{2,Coop}(P_{2,C})$. Then the induced welfare W_C is dominated by that achieved under the optimal IEA of type 1, i.e. $W_C^1 \geq W_C$.*

Proof: As shown in proof of proposition 13, the choice set for IEA type 1 can be expanded to a convex set with non-empty interior: $E_1 \geq 0, E_2 \geq 0$ and

$$u'(E_1) \leq D'(P_1 + nE_1) + \beta(1-\delta)D'\big((P_1 + nE_1)(1-\delta) + nE_2\big)$$

We will show that under the given conditions, the emissions under optimal IEA of the given type, $(E_{1,C}, E_{2,C})$, is contained in this choice set. This implies welfare obtained under optimal IEA type 1 cannot be inferior. Using the first-order necessary condition for E_1, given the announcement $E_{2,C}(P_2)$, we have

$$\begin{aligned} u'(E_{1,C}) &\leq D'(P_1 + nE_{1,C}) + \beta(1-\delta)D'\big(P_{3,C}\big) - \beta(1-\delta)E'_{2,C}(P_{2,C})\big(u'(E_{2,C}) - nD'(P_{3,C})\big) \\ &\leq D'(P_1 + nE_{1,C}) + \beta(1-\delta)D'\big(P_{3,C}\big) \end{aligned}$$

The inequality follows from (i) $E_{2,C} \geq E_{2,Coop}(P_{2,C})$, utility is concave and damage is convex, so that we have $u'(E_{2,C}) - nD'(P_{3,C}) \leq 0$, and (ii) $E'_{2,C}(P_2) \leq 0$. Therefore allocation $(E_{1,C}, E_{2,C})$ is contained in the choice set of optimal IEA type 1 problem.■

D Proofs for Section 5

Proposition 16. *If $E_{1,Coop}(P_1), E_{2,Coop}(P_{2,Coop}) > 0$ holds, for any $\varepsilon > 0$ sufficiently small, there exist $\theta_0, \theta_1 > 0$ such that the announcement $E_{2,C}(P_2; \varepsilon) = \max\{0, \theta_0 - \theta_1 P_2\}$ implements the emission levels $(E_{1,Coop}, \frac{E_{2,Coop}}{1+\varepsilon})$ as a symmetric equilibrium.*

Proof: Fix initial pollution level P_1. If $\lim_{E \downarrow 0} u'(E) = \infty$ holds, the result follows immediately from proposition 15, and, for $\varepsilon > 0$ sufficiently small, both θ_0 and θ_1 are positive.[31]

Assume for the rest of the proof that $u'(0) < \infty$. Define $V_i(\{E_{1j}, E_{2j}\}_{j=1}^n)$ be the level of ex ante *private* value of the country i when countries emit $\{E_{1j}, E_{2j}\}_{j=1}^n$. Social welfare, given this profile is defined as $W(.) = \sum_{j=1}^n V_j(.)$. Obviously $W(\{E_{1j}, E_{2j}\}) \leq W(E_{1,Coop}, E_{2,Coop}) \equiv W_{coop}$ for any feasible emissions profile. Imagine an auxillary cooperative problem where the planner chooses first period emissions E_1 subject to the constraint $E_2 = 0$ for each country. This problem is well-behaved and has a unique solution under assumption 1. Call the resulting welfare $\tilde{W}$. Clearly $\tilde{W} < W_{coop}$, since the auxillary problem imposes extra constraints relative to the first-best. The inequality is strict since there is a unique solution to the planner's unconstrained problem by proposition 3, and it features $E_{2,Coop} > 0$ by assumption. Then, there exists $\eta > 0$ such that $W_{coop} - \eta > \tilde{W}$. By continuity, there exists some $\bar{\varepsilon} > 0$ such that $W_C(\varepsilon) \equiv W(E_{1,Coop}, \frac{E_{2,Coop}}{1+\varepsilon}) > W_{coop} - \eta$ for all $\varepsilon \in (0, \bar{\varepsilon}]$. Define θ_0 and θ_1 as in proposition 15 for emissions profile $(E_{1,Coop}, \frac{E_{2,Coop}}{1+\varepsilon})$. For $\varepsilon > 0$ sufficiently small, both θ_1 and θ_0 are positive. Fix such an $\varepsilon > 0$.

Now we claim that the announcement $E_{2,C}(P_2) = \max\{0, \theta_0 - \theta_1 P_2\}$ supports $(E_{1,Coop}, \frac{E_{2,Coop}}{1+\varepsilon})$ as a symmetric equilibrium. Given the proof of proposition 15 and the remark right after that, it suffices to verify that no country would deviate in period 1, pollute more in period 1 to emit zero in period 2. Suppose every player except for country 1 plays $E_{1,Coop}$ so that $P_2 = (1-\delta)(P_1 + (n-1)E_{1,Coop} + E_1)$ effectively becomes the "choice variable" for country 1 since it is uniquely determined by E_1. Obviously, a deviation to any point where $E_{2,C}(P_2) = \theta_0 - \theta_1 P_2$ cannot be optimal by revealed preference (by proposition 15).

Consider a deviation where E_1 is sufficiently high, so that resulting P_2 is over the region in which $E_{2,C} = 0$. Let $E_1 = E_{1,Coop} + n\Delta$, $\Delta > 0$ be this deviation when all other players continue playing $E_{1,Coop}$. Observe that the resulting second period pollution level is $P_2 = (1-\delta)(P_1 + nE_{1,Coop} + \Delta) > P^*_{2,C}$. Moreover, since there is commitment, the second period emission equals $\bar{E}_2 \equiv 0$ for all countries. If this deviation is an improvement for country 1, we must have

$$nV_1(E_{1,Coop} + n\Delta, E_{1,Coop}, ..., \bar{E}_2, \bar{E}_2, ...) > W_C(\varepsilon)$$

Now reshuffle the emissions in period 1 and consider $\bar{E}_1 \equiv E_{1,Coop} + \Delta$ for all countries. Clearly, damage values and pollution stock P_2 remains intact, but social welfare improves because utility function is

[31] Observe that the denominator of (21) is positive, because $E_{2,Coop}$ makes the denominator zero and the given profile features lower emissions. Fully cooperative emissions solve $u'(E_{1,Coop}) - nD'(P_1 + nE_{1,Coop}) - n\beta(1-\delta)D'(P_{2,Coop} + nE_{2,Coop}) = 0$. Since $n > 1$, for sufficiently small $\varepsilon > 0$, the numerator of (21) is also positive.

concave, i.e. $nu(E_{1,Coop} + \Delta) \geq u(E_{1,Coop} + n\Delta) + (n-1)u(E_{1,Coop})$. Moreover, emissions in period 2 are identical since all countries play non-cooperatively and P_2 did not change. The social welfare satisfies

$$W(\bar{E}_1, \bar{E}_2) \geq nV_1(E_{1,Coop} + n\Delta, E_{1,Coop}, ..., \bar{E}_2, \bar{E}_2, ...) > W_C(\varepsilon).$$

Profile $(\bar{E}_1, \bar{E}_2)$ just constructed is a feasible emissions in the auxilary problem where the planner chooses first period emissions subject to $E_2 = 0$, hence $\tilde{W} \geq W(\bar{E}_1, \bar{E}_2)$. Since $W_C(\varepsilon) > \tilde{W}$ by construction, we obtain a contradiction. ■

Lemma 8. *If $E_{1,Coop}(P), E_{2,Coop}(P_{2,Coop}) > 0$ holds, for $\varepsilon > 0$ sufficiently small, there exist $\theta_0, \theta_1 > 0$ such that the announcement $E_{2,C}(P_2) = \max\{0, \min\{E_{2,NC}(P_2), \theta_0 - \theta_1 P_2\}\}$ implements the emission levels $(E_{1,Coop}, \frac{E_{2,Coop}}{1+\varepsilon})$ as a symmetric equilibrium.*

Proof: We follow similar steps as in the proof of proposition 16. Define private value $V_i(.)$ and social welfare $W(.)$ in the same way. Fix initial pollution level P_1, imagine an auxillary cooperative problem where planner chooses first period emissions E_1 subject to the constraint that countries will act *non-cooperatively* in period 2 and follow $E_{2,NC}(P_2)$. Let $\tilde{W}_1$ be the welfare level at an optimal solution to this problem. Similarly, consider a second auxillary problem where the planner chooses first period emissions E_1 subject to the constraint $E_2 = 0$ for each country, call the resulting welfare level $\tilde{W}_2$. Clearly $\tilde{W} \equiv \max\{\tilde{W}_1, \tilde{W}_2\} < W_{coop}$, therefore, there exists $\eta > 0$ such that $W_{coop} - \eta > \tilde{W}$. By continuity, there exists some $\bar{\varepsilon} > 0$ such that $W_C(\varepsilon) \equiv W(E_{1,Coop}, \frac{E_{2,Coop}}{1+\varepsilon}) > W_{coop} - \eta$ for all $\varepsilon \in (0, \bar{\varepsilon}]$. Fix such an $\varepsilon > 0$. Define θ_0, θ_1 as in proposition 16.

Now we claim that the announcement $E_{2,C}(P_2) = \max\{0, \min\{E_{2,NC}(P_2), \theta_0 - \theta_1 P_2\}\}$ implements the emission levels $(E_{1,Coop}, \frac{E_{2,Coop}}{1+\varepsilon})$ as an equilibrium so that $W_C(\varepsilon) > \tilde{W}$. Suppose every player except for country 1 plays $E_{1,Coop}$ so that $P_2 = (1-\delta)(P_1 + (n-1)E_{1,Coop} + E_1)$ effectively becomes the "choice variable" for country 1 since it is uniquely determined by E_1. Obviously, a deviation to any point in the steeper portion of $E_{2,C}$ cannot be optimal by revealed preference.

We focus on the two cases separately.

(i) The first period emissions is sufficiently small, so that resulting P_2 is on the non-cooperative portion in Panel C of figure 2. Let $E_1 = E_{1,Coop} - n\Delta$, $\Delta > 0$ be this deviation when all other players continue playing $E_{1,Coop}$. Observe that the resulting second period pollution level is $P_2 = (1-\delta)(P_1 + nE_{1,Coop} - \Delta) < P^*_{2,C}$. Moreover, since there is commitment, the second period emission equals $\bar{E}_2 \equiv E_{2,NC}(P_2)$ for all countries. If this deviation is an improvement for country 1, we must have

$$nV_1(E_{1,Coop} - n\Delta, E_{1,Coop}, ..., \bar{E}_2, \bar{E}_2, ...) > W_C(\varepsilon)$$

Now reshuffle the emissions in period 1 and consider $\bar{E}_1 \equiv E_{1,Coop} - \Delta$ for all countries. Clearly, damage values and pollution stock P_2 remains intact, but social welfare improves because utility function is concave, i.e. $nu(E_{1,Coop} - \Delta) \geq u(E_{1,Coop} - n\Delta) + (n-1)u(E_{1,Coop})$. Moreover, emissions in period

2 are identical since all countries play non-cooperatively and P_2 did not change. From a social welfare point of view, we have

$$W(\bar{E}_1, \bar{E}_2) \geq nV_1(E_{1,Coop} - n\Delta, E_{1,Coop}, ..., \bar{E}_2, \bar{E}_2, ...) > W_C(\varepsilon).$$

Profile $(\bar{E}_1, \bar{E}_2)$ is a feasible emissions profile for the first auxilary problem in which the planner chooses first period emissions subject to the constraint that countries play non-cooperatively in period 2, hence $\tilde{W} \geq W(\bar{E}_1, \bar{E}_2)$. Since $W_C(\varepsilon) > \tilde{W}$ by construction, we obtain a contradiction.

(ii) Deviation E_1 is sufficiently high so that second period emissions equal zero. This case is already covered in the proof of proposition 16 and the details are omitted.

We have shown that $\tilde{W} \geq W(\bar{E}_1, \bar{E}_2)$. Since $W_C(\varepsilon) > \tilde{W}$ by construction, we obtain a contradiction.

■