

# International Environmental Agreements with Cooperative Research & Development\*

Ross A. Mohr<sup>†</sup>

September 7, 2010

## Abstract

This paper builds on a standard, linear model of International Environmental Agreements by allowing countries to invest in abatement cost-reducing Research & Development (R&D). This paper also considers two different cooperation regimes: pollution abatement and R&D investment can either be provided independently with two separate agreements, or countries may choose to negotiate provision of both goods in a single, joint agreement. I show that when the joint treaty achieves a high enough level of participation, which implies a threshold amount of R&D investment, even non-signatories find it individually rational to abate pollution. That is, the resulting technology lowers the cost of pollution abatement enough so that the behavior of non-signatories tips toward the full cooperative outcome for pollution abatement, eliminating the incentive to free-ride. In this case, a joint agreement for cooperation on the environment and R&D increases pollution abatement and aggregate welfare.

Keywords: Global public goods, International agreements, Climate change, Research & development

JEL classification: H41, Q54, D62

---

\*I thank Charlie Kolstad, Cheng-Zhong Qin, Perry Shapiro, and Hani Mansour for helpful comments.

<sup>†</sup>Department of Economics, 2127 North Hall, University of California, Santa Barbara, CA 93106-9210; email: mohr@econ.ucsb.edu

# 1 Introduction

As a cooperative solution to reduce global greenhouse gas emissions, the Kyoto Protocol has proven to be severely inadequate. Barrett (2008) gives three main reasons for this failure. First, part of the problem is that an International Environmental Agreement (IEA) should have significant *participation*, but many of the world's highest emitters are not members of the Kyoto Protocol. There is also a *compliance* issue since many signatories have not actually complied with their emissions targets. Finally, an environmental treaty should be *environmentally effective*; that is, it should deliver greater emissions reductions than in the absence of cooperation.<sup>1</sup> However, the paramount reason why cooperation on a climate change solution has been difficult is that an IEA must be *self-enforcing*; that is, since participation is voluntary and an agreement cannot be enforced by a World Government, it must be in a country's best interest to join the IEA and to comply with its prescription.

One of the main reasons for both low participation and insufficient compliance is that the perceived costs of pollution abatement greatly outweigh the benefits (at least in the short-run). Thus, lowering abatement costs by funding research and development (R&D) of new technologies has recently gained a higher priority as part of the solution to reducing greenhouse gas emissions. However, *required* R&D funding has not yet been included in any IEA; rather, existing IEAs such as the Montreal Protocol, which phased-out the use of ozone-destroying chemicals, promote R&D cooperation among members and even encourage technology diffusion to non-members (Barrett 2003).<sup>2</sup> Furthermore, Barrett (2009) notes that after the negotiation of the UN Framework Convention on Climate Change in 1992, energy R&D investment actually decreased among those countries that became members of the Kyoto Protocol and has only recently (in 2002) began to increase. To put the problem of insufficient energy R&D in a more historical perspective, Newell (2008) observes that the

---

<sup>1</sup>The words *agreement* and *treaty* are used interchangeably.

<sup>2</sup>Heal and Tarui (2010) also provide the example of the Asia-Pacific Partnership on Clean Development & Climate, which is a non-binding agreement to promote pollution abatement and R&D cooperation, among other environmental goals. See <http://www.app.gov/>

amount of publicly-funded energy R&D of members of the International Energy Agency was, in real terms, approximately the same in 2006 as it was in 1974 and only amounted to 4 percent of those countries' overall public R&D investment.

In the period since the negotiation of the Framework Convention on Climate Change, there have been several efforts at cooperative R&D Agreements (RDA). For instance, the United States has bilateral Science and Technology (S&T) agreements with 37 countries.<sup>3</sup> However, these agreements mostly serve to coordinate activities and facilitate collaboration, as they typically do not set explicit funding goals or targets for the countries. This is also the case for the Carbon Sequestration Leadership Forum: it seeks to develop lower-cost technologies for carbon capture and sequestration through international collaboration, but does not impose any binding commitments on its members.<sup>4</sup> Not all attempts at cooperative R&D, however, only aim to coordinate actions. The Seventh Framework Programme (FP7) of the European Community (EC), adopted on December 18, 2006 with a total budget of 32,413 million Euro (with 1890 million Euro to environmental and climate change research), serves to fund and coordinate R&D projects from 2007 to 2013.<sup>5,6</sup>

Since pollution abatement and the knowledge spillovers from R&D investment are both public goods with free-rider incentives, there are two market failures and both goods will likely be under-provided (Jaffe, Newell, and Stavins 2005). Thus, I consider a model in which countries may cooperate on the provision of *both* goods and begin the analysis with the general question: Will cooperative investment in R&D to produce abatement cost-reducing technologies increase participation in the IEA and, thus, increase global pollution abatement and welfare?

---

<sup>3</sup>For a complete list of countries and dates of Entry-Into-Force, see <http://www.state.gov/g/oes/rls/fs/2009/115031.htm>

<sup>4</sup>For more information, see <http://www.cslforum.org/index.html>

<sup>5</sup>For a complete list of members, including additional bilateral S&T agreements for funding and coordination, see [http://ec.europa.eu/research/fp7/index\\_en.cfm](http://ec.europa.eu/research/fp7/index_en.cfm)

<sup>6</sup>Golombek and Hoel (2005) doubt the effectiveness of a RDA with funding commitments because the amount of R&D investment is difficult to monitor. However, Newell (2008) says that “with regard to energy, the International Energy Agency already collects annual data on public energy R&D spending by [member] countries, a process that could be adjusted if necessary to serve a more formal purpose” (pg. 24).

In analyzing this question, this paper makes two main contributions to the literature. The first contribution focuses on the treatment of R&D knowledge spillovers with respect to cooperation incentives. Some previous papers have assumed that knowledge spillovers from cooperative R&D can be completely restricted from non-signatories or at least treated as an imperfect club good (Carraro and Siniscalco 1997; Buchner *et al.* 2005). However, as Barrett notes, "International agreements routinely encourage cooperation in R&D. But in no case do they seek to deprive non-signatories of the fruits of this cooperation" (Barrett 2003, p. 309). Furthermore, it may actually be in a member country's best interest to share the new technology with a non-member if it leads to even more pollution abatement. Thus, I relax this club good assumption on R&D spillovers and assume that all countries have access to the resulting technology *without restriction*; that is, regardless of a country's membership status in the IEA, it may use the new abatement cost-reducing technology produced by R&D.

The second contribution of this paper is to consider two different treaty structures, which provides more insight into how the treaties change the incentives of the countries. In one treaty regime, cooperation on abatement and R&D investment are negotiated independently as two separate agreements, and in the second one members cooperatively provide both abatement and R&D in one joint agreement.<sup>7</sup> I assume that countries are symmetric, which implies identical *ex ante* expected payoffs; thus, all countries will *ex ante* prefer the cooperation regime that has the greatest expected payoff.

The first result of the model is that when abatement and R&D investment are provided by two separate agreements, R&D investment does not act directly as an IEA participation incentive, and the RDA is formed by a subset of IEA members. However, the new technology from R&D investment does allow members of the IEA to abate pollution at lower cost than before, which increases global welfare. Then, I show that the joint agreement, in which members agree to provide both abatement and R&D investment, achieves a higher level of

---

<sup>7</sup>These two different cooperation regimes are also analyzed in Carraro and Marchiori (2004). More details are provided in the following section.

R&D investment but a lower level of pollution abatement than if there were two separate agreements.

Finally, the main result of this paper is that when the jointly-negotiated agreement sustains a high enough participation rate to cross a certain threshold (which also implies a threshold level of R&D investment) the resulting technology lowers the cost of abatement enough such that the behavior of non-signatories tips, and it actually becomes individually rational for non-signatories to abate pollution. In this case, coordination is needed so that the threshold level of R&D investment is met. I show that the joint agreement emerges endogenously as the preferred cooperation regime and unambiguously increases aggregate welfare. Thus, cooperative R&D investment may be able to tip the behavior of non-signatories so that all countries find it in their best interest to abate pollution, eliminating the incentive to free-ride on pollution abatement.

## 2 Background

Early IEA research provides the pessimistic and paradoxical result that when the potential gains to cooperation are large, a self-enforcing IEA suffers from a low participation rate and cannot improve much over the non-cooperative level of pollution abatement; however, when the gains to cooperation are small, a self-enforcing IEA can sustain a high level of participation, but the outcome still will not be much better than the non-cooperative outcome.<sup>8</sup> Carraro and Siniscalco (1995, 1997) show that it is possible (theoretically, at least) to deter free-riding on the IEA and increase participation incentives by linking an IEA with R&D cooperation.<sup>9,10</sup> Continuing in this line of research, Buchner *et al.* (2005) analyze how linking the Kyoto Protocol with R&D cooperation may change the participation incentives of

---

<sup>8</sup>This result can be found in Barrett (1994), Carraro and Siniscalco (1993), and Hoel (1992). Barrett (2003) provides a comprehensive survey of the IEA literature.

<sup>9</sup>In one of the first analyses of issue linkage, Cesar and de Zeeuw (1996) show that payoff asymmetries in Prisoner's Dilemma games can be resolved efficiently by linkage.

<sup>10</sup>Although the main focus of this paper is IEA participation incentives and abatement provision, early theoretical results showing that cooperation can internalize R&D knowledge spillovers can be found in Katz (1986), d'Aspremont and Jacquemin (1988), and Poyago-Theotoky (1995).

the United States; however, they find that the linkage would not be *credible* and that the US would remain a non-signatory.<sup>11</sup>

A crucial assumption that drives the results of Carraro and Siniscalco (1997) and Buchner *et al.* (2005) is that R&D knowledge is a club good: IEA members share knowledge spillovers between themselves, but can completely restrict spillovers from non-signatories. The reason for this is that if the benefit from new abatement cost-reducing technologies is restricted to signatories only, then linking the IEA with R&D cooperation should increase participation incentives for the IEA. Treating R&D innovations as a club good may make the economics of cooperation on pollution abatement more favorable, but as noted in the introduction, it is not necessarily realistic.

Furthermore, Carraro and Marchiori (2004) analyze how the incentives to abate pollution change when countries can choose to form an agreement to cooperatively provide both abatement and R&D (a linked agreement), but still assuming that R&D is a club good. I also consider a joint treaty in this paper; however, Carraro and Marchiori (2004) assume that abatement, a pure public good, is more prone to free-riding than R&D and so R&D knowledge is again used as a participation incentive.

In general, R&D knowledge and new technology can spillover to other countries in several ways, including: selling the new technology to foreign firms and countries, reverse-engineering of a patented technology, or simply sharing knowledge. Newell (2008) also suggests that since patents are used to protect the value to the inventor, perhaps one goal of a RDA could be to purchase the new technology for public use (or award a prize to the inventor), thus allowing for even greater R&D spillovers while still securing intellectual property rights.

In nearly all papers that model both R&D investment and environmental policy, the main effect of the new technology is to lower total and marginal abatement costs.<sup>12</sup> Typically,

---

<sup>11</sup> *Credibility* means that the threat to exclude a country from R&D cooperation if it does not cooperate on the environmental agreement is binding. Tol, Lise, and van der Zwaan (2000) and Barrett (2003) also discuss the credibility of linked negotiations.

<sup>12</sup> Heal and Tarui (2010) also consider how new technologies can result in a completely different cost structure than the old technology; thus, total (and fixed) costs may fall, but a higher variable cost can *increase* the marginal abatement cost of achieving a certain level of abatement.

aggregate R&D is taken to be the sum of all R&D investments while taking into account spillovers from other countries or firms. Also common to the majority of this literature is that countries minimize a convex cost function that includes environmental damages, R&D investment, and abatement costs, which can be reduced through R&D, to find their optimal amounts of abatement and R&D. Since one of the goals of this paper is to derive analytical solutions for the sizes of different agreements, I begin with the standard, linear IEA model and then adopt a payoff function that incorporates abatement cost-reducing technologies in a very simple way: R&D investment reduces the cost of abatement by a proportion of the total amount invested.<sup>13</sup> A similar payoff function can be found in Barrett (2006).

In fact, the model in this paper most closely resembles that of Barrett (2006), who considers a treaty system to fund and promote adoption of a "breakthrough" technology.<sup>14,15</sup> He shows that if there are increasing returns to adoption in the breakthrough abatement technology, then there is a tipping point such that if the number of countries adopting the technology passes this amount, it becomes in the best interest of all other countries to adopt the technology. This paper does not consider a breakthrough technology, *per se*, but rather that the cost of pollution abatement can be reduced through R&D investment. Similarly to Barrett (2006), though, I examine a case in which R&D investment has a tipping point, but here it affects a country's pollution abatement incentives, not the incentives for more countries to invest in R&D.

Barrett (2009) gives a general overview of the technologies, both carbon-free and carbon-reducing, that may be included in a "climate-technology revolution" and discusses the viability of those technologies, the risks versus benefits, and other economic considerations. In this paper, the way that R&D investment is modeled is broadly compatible with resulting in

---

<sup>13</sup>A linear IEA model has been used recently by Barrett (2001 and 2003) and Kolstad (2007). The linear payoff function is an extremely simplified version of reality, but it does allow one to derive analytical results. Additionally, it does produce results that are consistent with more general functional forms (Barrett 2003).

<sup>14</sup>In general, a "breakthrough" technology can be thought of as a zero-emissions energy technology. Hoffert *et al.* (2002) survey a wide variety of energy technologies in terms of current limitations and potential future breakthroughs.

<sup>15</sup>Buchner and Carraro (2005) also consider a climate agreement based solely on R&D cooperation. In their model, R&D knowledge spillovers are an imperfect club good.

any new technology that reduces the abatement costs of existing, greenhouse gas-emitting technologies, and "breakthrough" carbon-free technologies, such as wind or solar, are only treated as the limiting case of fully-cooperative R&D investment. Abatement cost-reducing technologies that are consistent with my model include those that are complementary to existing polluting technologies - for example, a new technology to reduce the cost of carbon capture and sequestration or to reduce the cost of decarbonizing ("cleaning") fossil fuel before combustion (end-of-pipe technologies). Another example of a technology compatible with my model is a new, cleaner technology that substitutes for an existing one and, thus, allows a certain level of abatement to be achieved at lower cost.

## 3 Model

### 3.1 Benchmark IEA

Before introducing the R&D component of the model, I begin by deriving the properties of the benchmark, linear, self-enforcing IEA.<sup>16</sup> Let there be  $N \geq 3$  countries that each make a one-shot decision to either Abate or Pollute. Assume that all countries have *ex ante* identical payoff functions. Since pollution abatement is a global public good, the abatement of one country benefits all other countries. Let each country's payoff be a linear function of the benefit from total abatement and its own abatement cost. The payoff of country  $i \in N$  is:

$$\pi_i(q_i, Q) = b(q_i + Q_{-i}) - cq_i, \tag{1}$$

where  $q_i \in \{0, 1\}$  is the choice to Abate or Pollute ( $q_i = 1$  if country  $i$  plays Abate) and  $Q_{-i}$  is the total amount of abatement by all countries except country  $i$ , so that  $Q = \sum_{j=1}^N q_j = q_i + Q_{-i}$  is the total amount of pollution abatement. Note that since Abate takes the value  $q_i = 1$ , the total amount of pollution abatement,  $Q$ , is also the *number* of countries that abate pollution. The parameter  $b > 0$  denotes the marginal benefit of abatement (by any

---

<sup>16</sup>See Barrett (2003) for a more complete treatment of this model.

country), and  $c > 0$  is the marginal cost of abatement.

The principle characteristic of global pollution problems is that since all countries benefit from the abatement efforts of a single country, each country would prefer to free-ride on the efforts of others, and pollution abatement is under-provided. In the context of this linear, discrete choice model, Pollute ( $q_i = 0$ ) is a country's dominant strategy, and the Nash equilibrium of this problem is that all countries play Pollute.<sup>17</sup> This implies that

$$\pi_i(0, Q_{-i}) - \pi_i(1, 1 + Q_{-i}) > 0, \quad (2)$$

which reduces to  $c > b$ . Furthermore, since all countries act out of self-interest, the socially optimal outcome in which all countries play Abate will not be achieved, even though doing so would make all countries better-off. All countries are made better-off by playing Abate when

$$\sum_{i=1}^N \pi_i(1, N) = \sum_{i=1}^N [bN - c] > 0, \quad (3)$$

which implies that  $Nb > c$ . In words, the primitive characteristic of this simplified version of the global pollution problem is that the marginal cost of abatement is greater than the marginal *individual* benefit of abatement but is less than the marginal *aggregate* benefit of abatement. Finally, this also implies that the total gain from full cooperation,  $N(Nb - c)$ , is positive.

The size of the self-enforcing IEA is derived as the subgame perfect equilibrium of a two-stage game. In the first stage, the membership stage, all countries decide individually whether or not to be a member of the IEA. In the second stage, the abatement stage, members of the IEA collectively choose to either Abate or Pollute in order to maximize their aggregate payoff, and non-signatories of the IEA simultaneously and non-cooperatively choose to either Abate or Pollute.<sup>18</sup> Solving the game by backwards induction, in the second

---

<sup>17</sup>In this paper, I only consider pure strategies.

<sup>18</sup>In Barrett (1994), the IEA is modeled as a three-stage game in which signatories act as Stackelberg leaders. However, it is conventional with linear payoffs to model signatories and non-signatories acting simultaneously since the discrete nature of the problem precludes strategic reactions by non-signatories.

stage non-signatories will always play Pollute, their dominant strategy. Then, assume that  $k$  countries join the IEA in the first stage. These  $k$  countries will maximize their collective payoff in the second stage by playing Abate if

$$\pi_{i \in k}(1, k) = bk - c \geq 0, \quad (4)$$

which implies  $k \geq \frac{c}{b}$ . This condition ensures that each signatory is made better off by playing Abate. Finally, since the number of countries in the IEA is an integer, let  $k = k^*$  be the smallest integer such that  $k \geq \frac{c}{b}$ . Thus, the equilibrium size of the IEA formed in the first stage satisfies

$$\frac{c}{b} + 1 > k^* \geq \frac{c}{b}.^{19} \quad (5)$$

The resulting payoffs to members and non-members are

$$\pi_{i \in k} = bk^* - c, \text{ and} \quad (6)$$

$$\pi_{i \notin k} = bk^*, \text{ respectively,} \quad (7)$$

and the aggregate payoff of all countries is

$$\Pi_{IEA} = Nbk^* - k^*c. \quad (8)$$

In order for the IEA of size  $k^*$  to be *self-enforcing*, signatories cannot become better off by defecting from the treaty and playing Pollute, nor can non-signatories become better off by acceding to the treaty.<sup>20</sup> As a consequence of assuming symmetric countries and a linear payoff function, every country  $i \in k^*$  is pivotal, which implies that any defection will cause all other members to play Pollute resulting in all countries having a payoff of zero. Also, it is not rational for any non-signatory to accede to the agreement because members are

---

<sup>20</sup>In other words, self-enforcement requires that treaty membership be individually rational. This is derived from cartel stability (d'Aspremont, Jacquemin, Gabszewicz, and Weymark 1983).

required to play Abate and the abatement cost that the country would incur,  $c$ , is greater than the additional benefit,  $b$ .

Finally, I can summarize the main result of the basic IEA model. Note that the size of the IEA,  $k^*$ , is decreasing in  $b$ , but that the total gain from full cooperation,  $N(Nb - c)$ , is increasing in  $b$ . In words, when the potential gains from international cooperation on pollution abatement are large, fewer countries will cooperate.

### 3.2 Including R&D investment

Now, in addition to choosing Abate or Pollute, each country also simultaneously chooses to either invest in R&D or not, which is denoted by "R&D" and "No R&D," respectively. Since both abatement and R&D investment are public goods, all countries benefit equally from the abatement efforts of any country, and all countries *may* benefit from any country's R&D. However, only a country that abates pollution (and incurs the abatement cost) will directly benefit from R&D investment because the only benefit of R&D in this model is to lower the cost of pollution abatement.

As in the benchmark IEA model, each country's payoff is a linear function of its choices and the actions of all other countries. The payoff of country  $i$  is now:

$$\pi_i(q_i, x_i, Q, X) = bQ - c[1 - \varepsilon(x_i + X_{-i})]q_i - dx_i \quad (9)$$

where  $x_i \in \{0, 1\}$  is the choice to invest in R&D or not ( $x_i = 1$  if country  $i$  invests in R&D) and  $X_{-i}$  is the total amount of R&D investment by all countries except country  $i$ , so that  $X = \sum_{j=1}^N x_j = x_i + X_{-i}$  is the total amount of R&D investment (and also the *number* of countries investing in R&D). I assume that R&D knowledge is a pure public good with full spillovers. This is most plausible if the new technology from R&D, which I show later is always provided cooperatively, is "owned" by a RDA and then made available to any other countries "costlessly" and without restriction (Newell 2008). The parameter  $\varepsilon \in (\frac{d}{Nc}, \frac{1}{N})$

represents the *effectiveness* of R&D investment, which is interpreted as the rate at which the resulting technology from R&D investment reduces the abatement cost. The upper bound on  $\varepsilon$  is necessary so that as  $X$  approaches  $N$  (full cooperation in R&D investment, the abatement cost  $c(1 - \varepsilon X)$  stays positive. The lower bound on  $\varepsilon$ , as will be explained later in the paper, is necessary to bound the sizes of the agreements below  $N$ . Finally, the marginal cost of R&D investment is  $d$ , and the pollution abatement variables are defined as in the previous section.

The marginal cost term,  $c(1 - \varepsilon X)$ , embodies this paper's main assumptions for R&D investment. I assume that the total amount of R&D investment produces a new technology that reduces abatement costs by a proportion of the total amount invested, as discussed in the background literature section, and that this new technology is available to all countries without restriction.<sup>21</sup> Thus, for a total amount of R&D investment,  $X$ , the marginal cost of abatement *for all countries* shrinks by the proportion  $(1 - \varepsilon X)$ ; however, if the total amount of R&D investment is zero ( $X = 0$ ), then the model simply reverts to the standard IEA in the previous section.

Since this is an extension of the benchmark IEA model, the main properties of the global pollution problem still hold, and Pollute is still a dominant strategy for country  $i$ , no matter if country  $i$  invests in R&D or not and taking the actions of all others as given. Thus,

$$\pi_i(0, x_i, Q, X) - \pi_i(1, x_i, Q, X) > 0 \tag{10}$$

must hold, which implies that  $c(1 - \varepsilon X) > b$  and  $X \in [0, \frac{c-b}{c\varepsilon}]$ .<sup>22</sup> As in the benchmark case, the dominant strategy is Pollute because the marginal cost of abatement is greater than the marginal benefit of abatement. Since this is true in the absence of R&D ( $X = 0$ ), the inequality  $c > b$  still holds. The upper bound on  $X$  implies that if the total amount

---

<sup>21</sup>I also assume that there is no uncertainty: R&D investment produces a new technology with 100 percent probability. I also assume that there are no time lags in diffusing and adopting the new technology since this is a one-time decision. Finally, I assume that there is no cost to adopting and implementing the new technology.

<sup>22</sup>The upper bound on  $X$  is derived by rearranging the dominant strategy condition.

of R&D investment is high enough (greater than  $\frac{c-b}{c\varepsilon}$ ), then the resulting technology causes Pollute to no longer be a dominant strategy. That is, if abatement costs are reduced enough through R&D investment, then the primitives of the pollution problem change. Clearly, this restriction on  $X$  plays a large role in determining how countries cooperatively provide abatement and will be examined in more detail later in the paper. However, for this section of the paper and through Section 5, Pollute is the dominant strategy for the abatement choice, which implies that  $X \in [0, \frac{c-b}{c\varepsilon})$ ; in other words,  $X \in [0, \frac{c-b}{c\varepsilon})$  puts the focus on the case in which cooperation is needed most.

As noted in the introduction, R&D investment also suffers from a free-rider incentive (since any country would prefer to let another do the research) and is under-provided with respect to the fully-cooperative, social optimum. In the context of this discrete choice model, No R&D ( $x_i = 0$ ) is a country's dominant strategy for the choice to invest in R&D or not. The first implication is that the strategy {Pollute, No R&D} strictly dominates the strategy {Pollute, R&D}, which implies that  $d > 0$ . The interpretation is that since the only *direct* benefit of investing in R&D is a reduction in a country's own abatement cost, no country would ever choose to invest in R&D unless it is also choosing to abate pollution.<sup>23</sup>

The second, stronger implication of the free-rider incentive for R&D investment is even if a country has chosen to abate pollution, it would still prefer to free-ride on another country's R&D. Thus, No R&D is also a dominant strategy for country  $i$  even if it is playing Abate. This implies that

$$\pi_i(1, 0, Q, X) - \pi_i(1, 1, Q, X) > 0, \quad (11)$$

which reduces to  $d > c\varepsilon$ . In words, No R&D is a dominant strategy when the marginal cost of investing in R&D is greater than the marginal reduction in the abatement cost (the marginal benefit of R&D). Finally, in the non-cooperative Nash equilibrium, all countries play {Pollute, No R&D},  $Q = X = 0$ , and the payoff to each country is zero.

---

<sup>23</sup>If the RDA is not providing the new technology and countries or firms are selling it to others, then  $d$  could be interpreted as the "net cost" of investment. But in the context of the discrete choice model, under-provision still implies that  $x_i = 0$  is a dominant strategy, which implies  $d > 0$ .

For the combined problem of providing pollution abatement and R&D investment, the efficient, social optimum consists of all countries providing the two goods; in other words, the greatest aggregate payoff is achieved when all countries cooperate. The first implication of this is simply an extension of the discussion in the previous section: in the absence of any R&D investment, all countries would be better-off playing Abate. This implies that

$$\sum_{i=1}^N \pi_i(1, 0, N, 0) = \sum_{i=1}^N [bN - c] > 0, \quad (12)$$

which holds if  $Nb > c$ , as in the simple IEA case. Similarly, if the social optimum of R&D provision could be achieved, then the aggregate payoff would be greater than in the non-cooperative outcome. Since the only direct effect of R&D investment is to reduce one's own abatement costs, only countries that abate have any incentive of also providing R&D, and full cooperation on R&D will only be sustained when all countries are also abating pollution. Thus, the aggregate payoff when all countries play {Abate, R&D} (full cooperation) is greater than if all countries only played {Abate, No R&D} when

$$\sum_{i=1}^N \pi_i(1, 1, N, N) - \sum_{i=1}^N \pi_i(1, 0, N, 0) > 0, \quad (13)$$

which reduces to  $Nc\varepsilon > d$ .<sup>24</sup> In summary, the fundamental characteristics of the problem of providing abatement and R&D investment, which are both public goods and suffer from free-rider incentives, imply the following inequalities derived from the linear payoff function:

$$Nb > c > c(1 - \varepsilon X) > b \text{ for } X \in [0, \frac{c-b}{c\varepsilon}), \text{ and} \quad (14)$$

$$Nc\varepsilon > d > c\varepsilon. \quad (15)$$

---

<sup>24</sup>Additionally, the aggregate payoff for the efficient, full cooperative outcome ( $Q = X = N$ ) is greater than the aggregate payoff of the inefficient Nash equilibrium, which is true as long as  $N(b + c\varepsilon) > c + d$ . This condition, though, is redundant.

## 4 IEA with Cooperative R&D

### 4.1 Game Structure

Given the properties of the two public good provision problems and the conditions implied by the model, no single country has any incentive to supply either pollution abatement or R&D investment on its own, and the total provision of goods in the absence of cooperation is  $Q = X = 0$  (the Nash equilibrium). The only way that positive amounts of either good will be supplied is if a coalition forms to cooperatively provide the good. The goal of the rest of this paper is to analyze how cooperative R&D investment and the resulting abatement cost-reducing technology affect a country's incentive to cooperatively abate pollution.

This paper analyzes two different treaty structures, which will provide more insight into how treaties affect the incentives of the players in the game. In this model, countries may either cooperate on R&D independently of their IEA membership, or they may choose to negotiate R&D investment jointly with pollution abatement as a single, linked agreement. I assume that a country's membership in the IEA poses no additional requirements for that country's R&D investment, and vice versa.

In this paper, the countries play a one-shot, two-stage, non-cooperative game, which is similar to Carraro and Marchiori (2004), and the sizes of the agreements derive from a subgame perfect equilibrium. Before the proper two-stage game, there is a pre-stage in which each country decides non-cooperatively which treaty regime should be pursued. I assume that countries are symmetric, which implies identical *ex ante* expected payoffs; thus, all countries will *ex ante* prefer the cooperation regime that yields the greatest expected payoff.

In the pre-stage if countries choose to negotiate two separate treaties, then they play two, parallel treaty games. In the first stage, the membership stage, each country decides simultaneously whether to be a member of both the IEA and RDA, only the IEA or RDA, or to not cooperate at all. In the second stage, the provision stage, signatories of either

agreement simultaneously choose to collectively provide the good. That is, members of the IEA choose to abate if doing so maximizes their collective payoff, and members of the RDA choose to invest in R&D if doing so maximizes their collective payoff. I allow for there to be overlap in memberships: some members of the RDA may also be members of the IEA, and vice versa. Since this stage of the treaty game happens at the same time for both the IEA and the RDA, I abstract away from a timing issue: I assume that the new abatement cost-reducing technology resulting from the cooperative R&D investment of the RDA is available *at the same time* as when signatories of the IEA make their abatement decision. Thus, signatories of the IEA make their abatement decision taking into account the cost reduction from R&D investment. Simultaneously, in the provision stage non-members of any of the agreements non-cooperatively make their abatement and R&D investment decisions.

If in the pre-stage, however, countries agree to provide both pollution abatement and R&D investment in one agreement, then in the membership stage countries will choose whether or not to be a member of the joint agreement. In the provision stage, members of the joint agreement will choose to abate pollution *and* invest in R&D if this maximizes their collective payoff. Note that in the joint agreement, signatories cannot choose to provide only one good or the other: either both goods are provided or neither. Again, I abstract away from a timing issue: I assume that signatories of the joint agreement make their abatement decision taking into account the cost reduction from the new technology, even though technically abatement and R&D investment are provided *at the same time*. Simultaneously, in the provision stage non-members of the joint agreement non-cooperatively choose their abatement and R&D investment.

## 4.2 Two Separate Agreements: IEA and RDA

I begin with the case when in the pre-stage, countries choose to negotiate two separate agreements: the IEA and RDA. Recall that the provision problem for abatement and R&D investment in this model is summarized by conditions (14) and (15) and proceed by

backwards induction. In the provision stage, non-signatories of either agreement will play their dominant strategy of Pollute or No R&D. Also in the second stage, members of the IEA decide collectively whether to play Abate or Pollute, and members of the RDA decide collectively to play R&D or No R&D. For example, a country that is a member of the IEA, but not the RDA, would play No R&D in the provision stage. A country may be a member of both agreements. Recall that the only *direct* benefit of a country's R&D investment is to lower its own abatement cost. So if a country chooses to not abate pollution, then it has absolutely no incentive to invest in R&D, regardless of its membership status in either agreement. So if the RDA exists and has a positive number of members, then they must also all be members of the IEA (and play Abate). Let  $k_E$  be the number of members in the IEA,  $k_R$  be the number of members in the RDA, and  $k_{EO}$  be the number of members in the IEA *only*. Then,  $k_E = k_{EO} + k_R$ , and all members of the RDA are also members of the IEA.

Continuing with the provision stage, I derive conditions such that all members of the IEA play Abate and that all members of the RDA play R&D. A member of both agreements will play Abate and R&D if it cannot gain by leaving the RDA to only be a member of the IEA and if it cannot get a higher payoff by simply being a free-rider on both agreements. The condition that ensures that this member cannot gain by defecting from the RDA is

$$\pi_{i \in k_R}(1, 1, k_E, k_R) - \pi_{i \in k_{EO}}(1, 0, k_E, 0) \geq 0, \quad (16)$$

which implies that  $k_R \geq \frac{d}{c\varepsilon}$ .<sup>25</sup> Thus, as long as this condition holds, members of the RDA will play {Abate, R&D}. If a member of both agreements defected from the IEA, then all countries would play Pollute, there would be no gain to R&D investments, and so the RDA would also no longer provide R&D investment either. The condition that ensures that a

---

<sup>25</sup> $\pi_{i \in k_R}(1, 1, k_E, k_R) - \pi_{i \in k_{EO}}(1, 0, k_E, 0) = bk_E - c(1 - \varepsilon k_R) - d - (bk_E - c)$ . Note that any defection from the RDA causes all remaining members to play No R&D. Thus,  $X = 0$ , but there are still  $k_E$  countries that abate.

member of both agreements cannot gain by defecting from the IEA is

$$\pi_{i \in k_R}(1, 1, k_E, k_R) \geq 0, \quad (17)$$

which implies that  $k_E \geq \frac{1}{b}[c(1 - \varepsilon k_R) + d]$ . So, the number of countries in the IEA is a function of the amount of R&D investment,  $k_R$ . Finally, I address countries that only belong to the IEA and not to the RDA. A member of the IEA will play Abate if its payoff is at least as great as its payoff from non-cooperation. Thus,

$$\pi_{i \in k_{EO}}(1, 0, k_E, k_R) \geq 0, \quad (18)$$

which implies that  $k_E \geq \frac{c}{b}(1 - \varepsilon k_R)$ . However,  $k_E \geq \frac{c}{b}(1 - \varepsilon k_R)$  is not large enough to ensure that a member of both agreements has a positive payoff; thus,  $k_E \geq \frac{1}{b}[c(1 - \varepsilon k_R) + d]$  is the necessary condition.

The stage 1 equilibrium, when each country makes its membership decision, determines the equilibrium sizes of the agreements,  $k_E^*$  and  $k_R^*$ . As in the benchmark IEA case, define  $k_R = k_R^*$  as the smallest integer at least as large as  $\frac{d}{c\varepsilon}$  and define  $k_E = k_E^*$  as the smallest integer at least as large as  $\frac{1}{b}[c(1 - \varepsilon k_R) + d]$ .

**Proposition 1** *If the IEA and RDA are negotiated as two separate agreements and conditions (14) and (15) hold, then the equilibrium sizes of the IEA and RDA satisfy:*

$$\frac{1}{b}[c(1 - \varepsilon k_R^*) + d] + 1 > k_E^*(k_R^*) \geq \frac{1}{b}[c(1 - \varepsilon k_R^*) + d] \text{ and} \quad (19)$$

$$\frac{d}{c\varepsilon} + 1 > k_R^* \geq \frac{d}{c\varepsilon}, \text{ respectively.} \quad (20)$$

Note that condition (15) implies that  $\frac{d}{c\varepsilon} > 1$ , and comparing this with the above condition for  $k_R^*$  shows that the RDA has at least two members, when it exists. One can obtain the outermost bounds on  $k_E^*$  by substituting in the equilibrium necessary conditions for  $k_R^*$ . Using  $\frac{d}{c\varepsilon} + 1 > k_R^*$  yields the inequality  $k_E^* > \frac{c}{b}(1 - \varepsilon)$ , and using  $k_R^* \geq \frac{d}{c\varepsilon}$  gives the inequality

$\frac{c}{b} + 1 > k_E^*$ . Thus, the following bounds also hold:

$$\frac{c}{b} + 1 > k_E^* > \frac{c}{b}(1 - \varepsilon). \quad (21)$$

This condition shows that although the size of the IEA now depends on the amount of R&D investment, there is not necessarily a change in the amount of countries providing pollution abatement. In fact, there is *possibly* (though unlikely) now one less country in the IEA than in the benchmark case, which means that  $k^* \geq k_E^*$ . Also, note that when the RDA exists, there is now  $X = k_R^*$  amount of R&D investment, and condition (14) implies that  $k_R^* < \frac{c-b}{c\varepsilon}$ .

Since every country is pivotal, the two agreements are internally stable because a defection from either agreement would cause the remaining members to revert back to playing their dominant strategies (Pollute and No R&D), which would make the defecting country worse off. Additionally, it can be easily shown that both the IEA and the RDA are stable from accession (for a non-signatory, the marginal cost of joining an agreement is greater than the marginal benefit). Thus, the IEA and the RDA are self-enforcing.

By comparing the aggregate equilibrium payoff of this cooperation regime with the aggregate payoff of the benchmark IEA model, one can determine the welfare effect of including R&D investment with the potential for cooperative provision into the model. Members of both the IEA and RDA, members of just the IEA, and non-members of the IEA (free-riders) have the following equilibrium payoffs, respectively:

$$\pi_{i \in k_R} = bk_E^* - c(1 - \varepsilon k_R^*) - d, \quad (22)$$

$$\pi_{i \in k_{EO}} = bk_E^* - c(1 - \varepsilon k_R^*), \text{ and} \quad (23)$$

$$\pi_{i \notin k_E} = bk_E^*. \quad (24)$$

Comparing these payoffs, note that  $\pi_{i \notin k_E} > \pi_{i \in k_{EO}} > \pi_{i \in k_R} \geq 0$ , which means that free-riders on the IEA have the highest payoff while a member of both the IEA and the RDA has the

lowest payoff. The aggregate equilibrium payoff when there are two separate agreements is

$$\Pi_S = Nbk_E^* - k_E^*c(1 - \varepsilon k_R^*) - k_R^*d.^{26} \quad (25)$$

**Proposition 2** *When abatement cost-reducing R&D is provided by a RDA and provided that the size of the IEA is unchanged ( $k^* = k_E^*$ ), the aggregate equilibrium payoff to all countries increases.*

**Proof.** The statement is true if and only if  $\Pi_S \geq \Pi_{IEA}$ . After substituting and rearranging the inequality, the resulting inequality is

$$k_R^*(c\varepsilon k_E^* - d) \geq (Nb - c)(k^* - k_E^*). \quad (26)$$

The left-hand side is weakly positive since  $k_E^* \geq k_R^* \geq \frac{d}{c\varepsilon}$ . Recall that the discussion following condition (21) implied that it is *possible* (though unlikely) that  $k_E^*$  is smaller than  $k^*$  by 1. If this is the case, then the right-hand side is positive, and I am unable to derive an unambiguous result. However, if the size of the IEA is unchanged ( $k^* = k_E^*$ ), then the right-hand side equals 0. In this case,  $\Pi_S \geq \Pi_{IEA}$  holds. ■

In this model, including R&D investment does not increase the aggregate payoff, *per se*, because it is modeled as a discrete choice with an incentive to free-ride. However, when countries cooperate to form a RDA alongside an IEA, the gains from cheaper abatement outweigh the cost of R&D investment, and in aggregate, all countries are better off.

As with the benchmark IEA model, I will briefly discuss the static properties of  $k_R^*$ ,  $k_E^*$ , and the total gain to full cooperation,  $N[Nb - c(1 - \varepsilon N) - d]$ .<sup>27</sup> The basic result from the simple IEA case is that when the potential gains to cooperation are large, the size of the agreement will be small. Or, on the other hand, a high level of cooperation will only be sustained when it is not really needed (low potential gain to full cooperation). The

---

<sup>26</sup>  $k_R^*$  countries have the payoff  $\pi_{i \in k_R} = bk_E^* - c(1 - \varepsilon k_R^*) - d$ ,  $k_{EO}^*$  countries have the payoff  $\pi_{i \in k_{EO}} = bk_E^* - c(1 - \varepsilon k_R^*)$ , and  $(N - k_E^*)$  countries have the payoff  $\pi_{i \notin k_E} = bk_E^*$ .

<sup>27</sup>  $\sum_{i=1}^N \pi_i(1, 1, N, N) = N[Nb - c(1 - \varepsilon N) - d]$

same story holds with respect to the new IEA,  $k_E^*$ : it is decreasing in  $b$ , but the total gain to cooperation is increasing in  $b$ . Focusing on the marginal cost of R&D investment,  $d$ , though, gives an ambiguous result. A decrease in  $d$  increases the total gain to cooperation. But while a decrease in  $d$  lowers the amount of signatories to the RDA, it has an ambiguous effect on the size of the  $k_E^*$ . This is because both the cost of R&D investment,  $d$ , and the benefit of lower abatement costs,  $c\varepsilon k_R^*$ , play a role in determining  $k_E^*$ , but have opposite effects. So the inverse relationship between the size of an agreement and the gains to cooperation is mostly still intact.

### 4.3 Jointly-Negotiated Agreement

Now, I turn to the agreement that follows the pre-stage when countries prefer to provide both pollution abatement and R&D investment with a single, jointly-negotiated agreement and, as before, proceed by backwards induction. In stage 2, non-members of the agreement would choose to play {Pollute, No R&D} since that is the dominant strategy. To analyze provision by treaty members in the second stage, first assume that there are  $k_J$  members in the joint agreement. The  $k_J$  signatories will each play {Abate, R&D} if the individual payoff from doing so is at least as big as the payoff of playing {Pollute, No R&D}. Since the individual payoff to all players is zero when there is no abatement or R&D, the  $k_J$  signatories will each play {Abate, R&D} as long as

$$\pi_{i \in k_J}(1, 1, k_J, k_J) = bk_J - c(1 - \varepsilon k_J) - d \geq 0, \quad (27)$$

which holds for  $k_J \geq \frac{c+d}{b+c\varepsilon}$ . To solve the membership stage, define  $k_J = k_J^*$  as the smallest integer at least as large as  $\frac{c+d}{b+c\varepsilon}$ .

**Proposition 3** *If both pollution abatement and R&D investment are provided by a single, jointly-negotiated agreement and conditions (14) and (15) hold, then the equilibrium size of*

the agreement satisfies:

$$\frac{c+d}{b+c\varepsilon} + 1 > k_J^* \geq \frac{c+d}{b+c\varepsilon}. \quad (28)$$

As in the previously discussed agreements, any defection by a signatory causes all remaining signatories to revert back to their dominant strategy, in this case {Pollute, No R&D}, and all countries are made worse off. Also, if a non-signatory accedes to the agreement, then it must play {Abate, R&D}. It follows from conditions (14) and (15) that the marginal net benefit for a non-signatory to accede to the joint treaty is negative. The appropriate condition for this is

$$b - c[1 - \varepsilon(k_J^* + 1)] - d < 0, \quad (29)$$

which can be rearranged as

$$k_J^* < \frac{c-b}{c\varepsilon} + \frac{d}{c\varepsilon} - 1.^{28} \quad (30)$$

However, it is sufficient that  $X = k_J^* < \frac{c-b}{c\varepsilon}$  by condition (14).

Finally, the joint agreement has the same inverse relationship between its size and the potential benefits of full cooperation as do the previously discussed agreements. The net benefit of full cooperation,  $N[Nb - c(1 - \varepsilon N) - d]$ , is increasing in  $b$  and decreasing in  $d$ , but  $k_J^*$  is decreasing in  $b$  and increasing in  $d$ , respectively. The main difference between this comparison and that made for the case of two separate treaties is that a change in  $d$ , the marginal cost of R&D investment, no longer results in an ambiguous effect. With two separate treaties, an increase in  $d$  results in higher membership in the RDA, but has an ambiguous effect on the size of of IEA,  $k_E^*$ , and the amount of abatement. However, when negotiations are joined an increase in  $d$  clearly increases the size of the agreement,  $k_J^*$ , which results in higher aggregate abatement and R&D investment.

---

<sup>28</sup>By condition (15),  $\frac{d}{c\varepsilon} > 1$  which implies that  $\frac{c-b}{c\varepsilon} + \frac{d}{c\varepsilon} - 1 > \frac{c-b}{c\varepsilon}$ .

## 5 Comparative Statics and Pre-Stage

The main results of the model, as summarized in Propositions 1 and 3, have determined the equilibrium sizes of the three possible agreements, and I now show how they rank and compare welfare outcomes. Recall that since the choice of abatement and R&D investment are both discrete, the number of countries in a particular agreement is equal to the amount of the good being provided. Thus, in the case of two separate agreements,  $Q = k_E^*$  and  $k_R^* = X$ , and in the case of a single, joint agreement,  $Q = X = k_J^*$ .

**Proposition 4** *Let conditions (14) and (15) hold. Then,  $k_J^* \leq k_E^*$  if and only if  $k_R^* \leq \frac{c+d}{b+c\varepsilon}$ . Thus, the sizes of the three agreements satisfies:*

$$k_R^* \leq k_J^* \leq k_E^*. \quad (31)$$

**Proof.** Suppose that  $k_J^* \leq k_E^*$ . A necessary condition for the weak inequality  $k_J^* \leq k_E^*$  is that  $\frac{c+d}{b+c\varepsilon} \leq \frac{1}{b}[c(1 - \varepsilon k_R^*) + d]$ , which is just the lower bounds on the agreement sizes. Simplifying and rearranging results in  $k_R^* \leq \frac{c+d}{b+c\varepsilon}$ . But since  $\frac{c+d}{b+c\varepsilon} \leq k_J^*$  defines the size of the joint agreement, the sizes of the three agreements satisfies:  $k_R^* \leq k_J^* \leq k_E^*$ . ■

Note that  $k_R^* \leq k_J^*$  means that the joint agreement provides more R&D investment than the RDA, which implies a greater reduction in abatement costs; however,  $k_R^* \leq k_J^*$  also means that, in terms of the effect on aggregate payoff, more countries incur the cost of R&D investment. Also,  $k_J^* \leq k_E^*$  means that the joint agreement provides less pollution abatement than the IEA. Thus, this proposition implies that if the countries choose to negotiate provision of the two goods jointly as one agreement, then in equilibrium there will be less pollution abatement, but more R&D, than if the two agreements were negotiated separately. In other words, in the joint agreement there is less abatement, but each unit of abatement costs less due to the higher amount of R&D. On the other hand, when there are two separate agreements more countries abate pollution than would in the joint agreement, but they do so at a higher per-unit cost. Recall that since  $k_R^* \leq k_E^*$ , some countries in the

IEA may still be free-riding on the RDA; the joint agreement eliminates this possibility since all members incur both the abatement and investment costs. Thus, the intuition behind this result is that the joint agreement represents a trade-off in benefits and costs with respect to the two, separate treaties, which implies that the size of the joint agreement lies in between the IEA and RDA. However, if countries only care about pollution abatement, then clearly the joint agreement would leave all countries worse off than under the IEA and RDA.<sup>29,30</sup>

To determine the outcome on economic welfare, I compare the aggregate equilibrium payoffs between the two different cooperation regimes. If negotiations are joined in a single agreement, then the equilibrium payoffs for signatories and non-signatories are the following:

$$\pi_{i \in k_J} = bk_J^* - c(1 - \varepsilon k_J^*) - d, \text{ and} \quad (32)$$

$$\pi_{i \notin k_J} = bk_J^*, \quad (33)$$

which results in the aggregate equilibrium payoff

$$\Pi_J = Nbk_J^* - k_J^*[c(1 - \varepsilon k_J^*) - d]. \quad (34)$$

The jointly-negotiated agreement is a welfare improvement on the two, separate treaties if  $\Pi_J \geq \Pi_S$ . To aid comparison, substituting equations (25) and (34) gives

$$Nbk_J^* - k_J^*[c(1 - \varepsilon k_J^*) - d] \geq Nbk_E^* - k_E^*c(1 - \varepsilon k_R^*) - k_R^*d. \quad (35)$$

Then, rearranging and grouping terms yields

$$c[k_E^*(1 - \varepsilon k_R^*) - k_J^*(1 - \varepsilon k_J^*)] \geq Nb(k_E^* - k_J^*) + d(k_J^* - k_R^*), \quad (36)$$

---

<sup>29</sup>Carraro and Marchiori (2004) derive a similar result that the size of the joint agreement is between the IEA and RDA. However, in their analysis they begin with the assumption that the RDA is larger than the IEA (it is assumed that there is less free-riding on the RDA because R&D is a club good), so their ranking of agreement sizes goes in the reverse order.

<sup>30</sup>Since  $N \geq k_R$ , it is true that  $N \geq \frac{d}{c\varepsilon}$ . Rearranging gives  $\varepsilon \geq \frac{d}{Nc}$ , which is the lower bound given in the earlier section.

and all terms in regular parenthesis are positive. The term on the left-hand side is the total abatement cost savings from the joint agreement versus the IEA and RDA. It highlights the trade-off of more abatement being provided with less cost reduction under two, separate agreements (the term  $k_E^*(1 - \varepsilon k_R^*)$ ) compared with less abatement being provided with more cost reduction in the joint agreement (the term  $k_J^*(1 - \varepsilon k_J^*)$ ). I am unable to sign the entire bracketed term on the left-hand side, and it is positive or negative depending on parameter values. The first term on the right-hand side is the total net benefit of abatement of the IEA over the joint agreement, and the second term on the right-hand side is the total R&D cost savings of the RDA over the joint agreement. Since the entire right-hand side is positive, though, one cannot determine analytically if  $\Pi_J \geq \Pi_S$ .<sup>31</sup> However, a special case yields the following result:

**Proposition 5** *If  $k_R^* \leq k_J = k_E^*$  and conditions (14) and (15) hold, then  $\Pi_J \geq \Pi_S$ .*

**Proof.** Substituting  $k_J = k_E^*$  in condition (36) yields  $c\varepsilon k_J^*(k_J^* - k_R^*) \geq d(k_J^* - k_R^*)$ , which is true since  $k_J^* \geq k_R^* \geq \frac{d}{c\varepsilon}$  ■

Since Proposition 4 only ensures the weak inequality  $k_J \leq k_E^*$ , it is certainly a possibility that  $k_J = k_E^*$ . In this case, the number of countries providing abatement (and incurring the abatement cost) is same in either cooperation regime, which implies that the welfare comparison reduces to a comparison of the net benefit of R&D investment in the joint agreement versus the RDA. Since in this case the joint agreement produces more R&D investment and reduces abatement costs by more than the RDA, the joint agreement yields a higher aggregate payoff than the two separate agreements.

The discussion in this section can now be extended to examine the pre-stage. In the pre-stage, countries non-cooperatively choose between the two treaty regimes. However, since the subgame perfect equilibria are not unique, a country does not "know" if it will be a signatory or non-signatory at the beginning of the game - a country only knows its

---

<sup>31</sup>Simulating the payoffs with different parameter values does not show any clear relationship or yield any conclusions about possible conditions where one payoff is always greater than the other.

possible equilibrium payoffs depending on its potential memberships. Furthermore, without imposing any further preferences for abatement or free-riding on countries, one cannot say, for example, that a country that would be a member of both the IEA and RDA under separate agreements would also choose to be a member of the joint agreement. The only information available to a country at the preliminary voting stage is its expected equilibrium payoff under joint negotiations and its expected payoff under two separate agreements. Thus, to determine the outcome of the pre-stage, I compare the expected payoffs.<sup>32</sup>

The expected equilibrium payoff of a country under two separate agreements is

$$E(\pi_i|S) = \frac{k_R^*}{N} \pi_{i \in k_R} + \frac{(k_E^* - k_R^*)}{N} \pi_{i \in k_{EO}} + \frac{(N - k_E^*)}{N} \pi_{i \notin k_E}, \quad (37)$$

which after simplifying equals  $\frac{\Pi_S}{N}$ . Similarly, the expected equilibrium payoff of a country under the joint agreement can be shown to equal  $\frac{\Pi_J}{N}$ . Naturally, a comparison of the expected payoffs at this point results in the same inability to draw an unambiguous conclusion as does comparing aggregate payoffs. All that can be concluded is that whichever cooperation regime results in the greater expected equilibrium payoff for each country will be the one that all countries prefer, and as before, this conclusion depends on parameter values. However, the special case considered in Proposition 5 can be extended to the following result:

**Corollary 1** *If  $k_R^* \leq k_J^* = k_E^*$  and conditions (14) and (15) hold, then  $\Pi_J \geq \Pi_S$ . Thus, the jointly-negotiated agreement weakly increases aggregate welfare over the two separate agreements and is chosen in the pre-stage as the preferred cooperation regime.*

**Corollary 2** *Additionally, if  $k_R^* = k_J^* = k_E^*$ , then  $\Pi_J = \Pi_S$  and all countries are indifferent between the two cooperation regimes.*

With the exception these two special cases, I am unable to draw any further conclusions regarding when a certain cooperation regime will be preferred.

---

<sup>32</sup>This method of comparing *ex ante* expected payoffs is also used in Barrett (2002) and Finus and Maus (2008).

Up to this point, this paper has modeled the provision problem of two public goods in a simple way: countries prefer to free-ride on both goods, which leads to a non-cooperative Nash equilibrium with no provision ( $Q = X = 0$ ). In the context of the discrete choice model, the properties of the problem implied conditions (14) and (15), which included the restriction on R&D investment,  $X \in [0, \frac{c-b}{c\varepsilon})$ . Written in terms of the sizes of the RDA and joint agreement, this is equivalent to  $k_R^* \leq k_J^* < \frac{c-b}{c\varepsilon}$ . Even though this restriction, which depends on parameter values, is a bit artificial, it has been used to focus on the case in which cooperation on abatement is needed most. In the next section, I reconsider this condition and analyze how this changes the incentives to cooperate.

## 6 Tipping Non-signatory Behavior Through R&D

Although the fundamentals of the provision problem remain intact (free-rider incentives result in a non-cooperation), I now consider the possibility that the joint agreement is large enough and produces a sufficient amount of R&D investment such that Pollute no longer strictly dominates Abate for non-members. Hence, conditional on the level of R&D investment passing the tipping point, the abatement game has two Nash equilibria: all countries play Pollute or all countries play Abate. To achieve the fully cooperative outcome for abatement, coordination is needed to establish the joint agreement of size  $k_J^* \geq \frac{c-b}{c\varepsilon}$ , and the amount of R&D  $X = \frac{c-b}{c\varepsilon}$  serves as a tipping point.<sup>33</sup> Furthermore, since  $k_R^* \leq k_J^*$ , the joint agreement achieves the R&D investment threshold for a larger range of parameters than the RDA.

As in the earlier derivation of the equilibrium sizes of the agreements, the joint agreement must be internally stable, which is true as long as the net marginal benefit of joining the joint agreement for a non-signatory is negative (see condition (30)). In order to analyze the effects of only the joint treaty reaching the R&D tipping point, I assume that equilibrium

---

<sup>33</sup>There is a difference between full *cooperation* and full *participation*. As I show later in this section, full cooperation is possible (all countries abate) without full participation (not all countries belong to the joint agreement).

size of the RDA still satisfies  $k_R^* < \frac{c-b}{c\varepsilon}$ . Thus, the following condition summarizes the new restrictions on the amount of R&D investment (and the sizes of the agreements):

$$k_R^* < \frac{c-b}{c\varepsilon} \leq k_J^* < \frac{c-b}{c\varepsilon} + \frac{d}{c\varepsilon} - 1. \quad (38)$$

In words, even though the benefit of pollution abatement now outweighs the cost, the marginal cost of investing in R&D,  $d$ , is still too high to make accession to the joint treaty profitable (for a non-member). Since condition (14) is no longer completely accurate, it is now replaced by

$$Nb > c > c(1 - \varepsilon k_R^*) > b > c(1 - \varepsilon k_J^*), \quad (39)$$

where  $k_R^*$  and  $k_J^*$  satisfy condition (38). This ensures that only the joint agreement, and not the RDA, tips the abatement decision of non-signatories. The main difference between conditions (14) and (39) is the inequality  $b > c(1 - \varepsilon k_J^*)$ , which means that the marginal benefit of abating pollution is greater than the marginal cost *as long as* the joint agreement exists. The condition given in (15) is still valid. Now, I reconsider the equilibrium agreement sizes and aggregate payoffs with the new conditions and game structure.

First of all, there is no change in the *relative* sizes of the coalitions, and Proposition 4 ( $k_R^* < k_J^* \leq k_E^*$ ) still holds; note that the only change is the strict inequality between  $k_R^*$  and  $k_J^*$ , which is due to condition (38). However, the main difference in the model now is the actions of non-members of the joint agreement. Previously, these non-members would completely free-ride on the joint agreement and play {Pollute, No R&D}. Now that the joint agreement produces enough R&D to pass the tipping point, condition (39) summarizes the incentives of the underlying abatement game and Pollute no longer strictly dominates Abate for non-members. The implication is that in the equilibrium of the joint cooperation regime there are still  $k_J^*$  members of the joint agreement that play {Abate, R&D}, but now the non-signatory countries also play Abate. And as long as  $k_J^* < \frac{c-b}{c\varepsilon} + \frac{d}{c\varepsilon} - 1$ , non-signatory countries still do not invest in R&D. Thus, the joint agreement still produces more R&D

investment than the RDA, but now there is also more overall pollution abatement when negotiations are joined than with the IEA.

With all  $N$  countries, both signatories *and* non-signatories, abating pollution under the joint agreement, the equilibrium payoffs become

$$\pi'_{i \in k_J} = Nb - c(1 - \varepsilon k_J^*) - d, \text{ and} \quad (40)$$

$$\pi'_{i \notin k_J} = Nb - c(1 - \varepsilon k_J^*). \quad (41)$$

The aggregate equilibrium payoff of the joint agreement is now

$$\Pi'_J = N^2b - Nc(1 - \varepsilon k_J^*) - k_J^*d.^{34} \quad (42)$$

**Proposition 6** *Let conditions (15), (38), and (39) hold. Then,  $\Pi'_J \geq \Pi_S$ .*

**Proof.** Following the discussion in the previous section, the jointly-negotiated agreement is a welfare improvement on the status quo if  $\Pi'_J > \Pi_S$ . Substituting from (25) and (42) gives

$$N^2b - Nc(1 - \varepsilon k_J^*) - k_J^*d > Nb k_E^* - k_E^*c(1 - \varepsilon k_R^*) - k_R^*d. \quad (43)$$

Rearranging and grouping terms yields

$$(Nb - c)(N - k_E^*) + k_J^*(Nc\varepsilon - d) > k_R^*(c\varepsilon k_E^* - d). \quad (44)$$

Since  $N \geq k_E^*$  and conditions (15) and (39) ensure  $Nc\varepsilon > d$  and  $Nb > c$ , the left-hand side is strictly positive. To complete the proof, it suffices to show that  $k_J^*(Nc\varepsilon - d) > k_R^*(c\varepsilon k_E^* - d)$ . This is true because  $k_J^*(Nc\varepsilon - d) > k_R^*(Nc\varepsilon - d) \geq k_R^*(c\varepsilon k_E^* - d)$ , where  $k_E^* \leq N$ . This proves that  $\Pi'_J > \Pi_S$ . ■

As in the previous section, it is a simple extension to determine which cooperation regime

---

<sup>34</sup> $k_J^*$  countries have the payoff  $\pi_{i \in k_J}$ , and  $(N - k_J^*)$  countries have the payoff  $\pi_{i \notin k_J}$ .

will be preferred by all countries. All that one needs to do is compare the *ex ante* expected equilibrium payoffs between the two cooperation regimes.

**Corollary 3** *Let conditions (15), (38), and (39) hold. Then, the jointly-negotiated agreement strictly increases aggregate welfare over the two separate agreements and is chosen in the pre-stage as the preferred cooperation regime.*

**Proof.** The expected equilibrium payoff of a country under two separate agreements is given in equation (37), and the expected payoff for a country under joint negotiations is  $E(\pi_i|J) = \frac{k_J^*}{N}\pi_{i \in k_J} + \frac{(N-k_J^*)}{N}\pi_{i \notin k_J}$ , which after simplifying equals  $\frac{\Pi'_J}{N}$ . Since  $\frac{\Pi'_J}{N} > \frac{\Pi_S}{N}$ , all countries prefer the jointly-negotiated agreement, and aggregate welfare unambiguously increases. ■

These two results show that when the joint agreement sustains a level of participation high enough so that the level of R&D investment passes the tipping point, aggregate welfare strictly increases. Thus, although only the members of the joint agreement invest in R&D, all countries now abate pollution: signatories of the joint agreement abate pollution because doing so maximizes their joint payoff, and non-signatories abate pollution because that is their dominant strategy. Furthermore, even though the joint agreement does not sustain full participation, full cooperation on abatement is achieved. What enables this fully cooperative outcome is that the joint agreement is large enough such that the level of R&D investment and, hence, abatement cost reduction are enough to tip non-signatory behavior. In this case, it is in all countries' best interest to provide both pollution abatement and R&D investment with a single agreement. Thus, a high enough level of cooperative R&D investment can push the abatement strategies of non-signatories over a tipping point and eliminate the incentive to free-ride on pollution abatement.<sup>35</sup>

The results of this section, however, are not quite as positive as they may seem. Recall

---

<sup>35</sup>For even higher values of  $d$ ,  $k_R^* > \frac{c-b}{c\varepsilon}$  is possible, and the RDA tips the behaviors of non-signatories under the two separate agreements. However, this occurs for an even smaller range of parameters than for  $k_J^*$ , and it can be shown that in this case the joint agreement still emerges as the preferred cooperation regime.

that the potential gain to full cooperation,  $N[Nb - c(1 - \varepsilon N) - d]$ , is decreasing in  $d$ , but that the size of the joint agreement is increasing in  $d$ . Furthermore, higher levels of  $d$ , keeping  $N$ ,  $b$ ,  $c$  and  $\varepsilon$  constant, is what enables the joint agreement to pass the threshold level of R&D investment. Thus, for higher levels of the cost of investment, participation in the joint agreement is greater, but the gains to cooperation fall; and when the cost of investment is high enough so that the tipping point is reached, the gains to cooperation are even less than before.

## 7 Conclusion

Faced with the failure of the Kyoto Protocol to reduce emissions of greenhouse gases, countries are realizing that the way forward will require reducing the costs of pollution abatement. However, this introduces the problem of providing another public good, R&D investment. As noted in the introduction, nearly all cooperative R&D agreements serve only to facilitate knowledge spillovers and most do not explicitly require funding by member countries. Thus, this paper considers the problem of cooperatively providing two public goods, pollution abatement and R&D investment, which both suffer from free-riding and under-provision. To try to gain more insight into the incentives to cooperate, I analyze two different cooperation regimes: forming a separate IEA and RDA or forming a single, joint agreement to provide both goods.

I model the provision of abatement and R&D investment in a linear, discrete choice model, and I assume that the total amount of R&D investment produces a new technology that reduces abatement costs by a proportion of the total amount invested and that this new technology is available to all countries costlessly and without restriction. I find that even though the aggregate payoff increases when countries may cooperatively provide R&D investment with the RDA, I cannot show definitively which cooperation regime is preferred in the equilibrium. This is because of the tradeoff inherent in the joint agreement: it provides

less abatement than the IEA, but at lower cost since it provides more R&D than the RDA. Finally, I consider the special case in which the joint treaty invests in a sufficient amount of R&D such that the new technology causes Pollute to no longer strictly dominate Abate for non-signatories. In this case, I show that the signatories of the joint agreement still provide abatement and R&D investment, and non-signatories only free-ride on R&D - all countries now provide abatement. Thus, the joint treaty provides a strict increase in welfare over two separate treaties and emerges endogenously as the unanimously preferred treaty structure.

The IEA literature has a history of pessimistic results, and I admittedly do not provide very positive results even considering the special case with the R&D tipping point. Clearly, the tipping point is a special feature of this linear payoff function and will not arise in a model with strictly convex cost functions. Furthermore, there are numerous other unrealistic simplifications including: *ex ante* symmetric countries in both abatement and R&D (which implies that all signatories are pivotal), no room for strategic reaction by non-signatories, identical spillovers across countries, no uncertainty in R&D innovation, and no time lags in developing, diffusing, and adopting the new technology. These simplifying assumptions are necessary for this paper's analysis because it is written, in part, as a complement to the many simulation papers in the current IEA literature<sup>36</sup>, and one of the goals of this paper is to provide transparent, analytic results wherever possible.

However, the main results are still suggestive of possible future cooperation outcomes. For instance, there has not yet been any attempt at an agreement which requires signatories to commit to both binding abatement *and* R&D investment targets. In the context of this paper's simple model, a joint agreement is better at reducing free-riding on R&D than is the RDA and with potentially only a small reduction in abatement with respect to the IEA (Proposition 4). The implication for future climate treaty negotiations is that even though R&D investment is important for lowering abatement costs, a joint treaty may be needed to keep participation incentives all pointed in the same direction.

---

<sup>36</sup>For instance, Carraro and Siniscalco (1997), Buchner *et al.* (2005), and Buchner and Carraro (2005).

Another suggestive point, which is discussed by Barrett (2006), is that cooperative R&D projects should be chosen strategically; thus, cooperative R&D projects for technologies that may encourage free-riders to abate at more socially-optimal levels should be pursued, even if there may be *better* technologies possible.<sup>37</sup> In this case, a joint agreement to provide both abatement and R&D investment would provide a better aggregate outcome, and cooperative R&D projects should focus on technologies that are more likely to be diffused and adopted by all countries, which would encourage lower levels of emissions among even non-cooperating countries.

## 8 References

Aldy, Joseph. E., Scott Barrett, and Robert N. Stavins. 2003. "Thirteen Plus One: A Comparison of Global Climate Policy Architectures." *Climate Policy*, 3: 373-397.

Barrett, Scott. 1994. "Self-Enforcing International Environmental Agreements." *Oxford Economic Papers*, 46: 878-894.

Barrett, Scott. 2001. "International Cooperation for Sale." *European Economic Review*, 45: 1835-1850.

Barrett, Scott. 2002. "Consensus Treaties." *Journal of Institutional and Theoretical Economics*, 158: 529-547.

Barrett, Scott. 2003. *Environment & Statecraft*. New York: Oxford University Press.

Barrett, Scott. 2006. "Climate Treaties and 'Breakthrough' Technologies." *The American Economic Review*, 92(2): 22-25.

Barrett, Scott. 2008. "Climate Treaties and the Imperative of Enforcement." *Oxford Review of Economic Policy*, 24(2): 239-258.

Barrett, Scott. 2009. "The Coming Global Climate-Technology Revolution." *The Journal of Economic Perspectives*, 23(2): 53-75.

---

<sup>37</sup>For completeness, I again note that my model is very similar to Barrett (2006), although he only considers cooperation on R&D investment.

Buchner, Barbara, and Carlo Carraro. 2005. "Economic and Environmental Effectiveness of a Technology-based Climate Protocol." *Climate Policy*, 4: 229-248.

Buchner, Barbara, Carlo Carraro, Igor Cersosimo, and Carmen Marchiori. 2005. "Back to Kyoto? US Participation and the Linkage between R&D and Climate Cooperation." In *The Coupling of Climate and Economic Dynamics*, ed. A. Haurie and L. Vignuiet. Dordrecht, The Netherlands: Springer.

Carraro, Carlo, and Carmen Marchiori. 2004. "Endogenous Strategic Issue Linkage in International Negotiations." In *Game Practice and the Environment*, ed. C. Carraro and V. Fragnelli. Cheltenham: Edward Elgar.

Carraro, Carlo, and Domenico Siniscalco. 1993. "Strategies for the International Protection of the Environment." *Journal of Public Economics*, 52: 309-328.

Carraro, Carlo and Domenico Siniscalco. 1995. "Policy Coordination for Sustainability: Commitments, Transfers, and Linked Negotiations." In *The Economics of Sustainable Development*, ed. I. Goldin and L. A. Winters. Cambridge: Cambridge University Press.

Carraro, Carlo and Domenico Siniscalco. 1997. "R&D Cooperation and the Stability of International Environmental Agreements." In *International Environmental Negotiations*, ed. C. Carraro. Cheltenham: Edward Elgar.

Cesar, Herman and Aart de Zeeuw. 1996. "Issue Linkage in Global Environmental Problems." In *Economic Policy for the Environment and Natural Resources*, ed. A. Xepapadeas. Cheltenham: Edward Elgar.

d'Aspremont, Claude and Alexis Jacquemin. 1988. "Cooperative and Noncooperative R&D in Duopoly with Spillovers." *The American Economic Review*, 78(5): 1133-1137.

d'Aspremont, Claude, Alexis Jacquemin, Jean Jaskold Gabszewicz, and John A. Weymark. 1983. "On the Stability of Collusive Price Leadership." *Canadian Journal of Economics*, 16: 17-25.

Finus, Michael and Stefan Maus. 2008. "Modesty May Pay!" *Journal of Public Economic Theory*, 10(5): 801-826.

- Golombek, Rolf and Michael Hoel. 2005. "Climate Policy under Technology Spillovers." *Environmental & Resource Economics*, 31: 201-227.
- Heal, Geoffrey and Nori Tarui. 2010. "Investment and Emission Control Under Technology and Pollution Externalities." *Resource and Energy Economics*, 32: 1-14.
- Hoel, Michael. 1992. "International Environment Conventions: The Case of Uniform Reductions of Emissions." *Environmental and Resource Economics*, 2: 141-159.
- Hoffert, Martin I., *et al.* 2002. "Advanced Technology Paths to Global Climate Stability: Energy for a Greenhouse Planet." *Science*, 298: 981-987.
- Jaffe, Adam B., Richard G. Newell, and Robert N. Stavins. 2005. "A Tale of Two Market Failures: Technology and Environmental Policy." *Ecological Economics*, 54: 164-174.
- Katz, Michael L. 1986. "An Analysis of Cooperative Research and Development." *Rand Journal of Economics*, 17(4): 527-543.
- Kolstad, Charles D. 2007. "Systematic Uncertainty in Self-enforcing International Environmental Agreements." *Journal of Environmental Economics and Management*, 53: 68-79.
- Newell, Richard G. 2008. "International Climate Technology Strategies." Discussion Paper 2008-12. Cambridge, Mass.: Harvard Project on International Climate Agreements, October 2008.
- Poyago-Theotoky, Joanna. 1995. "Equilibrium and Optimal Size of a Research Joint Venture in an Oligopoly with Spillovers." *The Journal of Industrial Economics*, 43(2): 209-226.
- Tol, Richard S. J., Wietze Lise, and Bob van der Zwaan. 2000. "Technology Diffusion and the Stability of Climate Coalitions." FEEM Working Paper No. 20.2000. Available at SSRN: <http://ssrn.com/abstract=229260>