The Dynamics of Climate Agreements[†]

Bård Harstad[‡]

6 July 2010

Abstract

This paper provides a model in which countries over time pollute as well as invest in technologies (renewable energy sources or abatement technologies). Without a climate treaty, the countries pollute too much and invest too little, partly to induce the others to pollute less and invest more in the future. Nevertheless, short-term agreements on emission levels can reduce welfare, since countries invest less when they anticipate future negotiations. The optimal agreement is tougher and more long-term if intellectual property rights are weak. If the climate agreement happens to be short-term or absent, intellectual property rights should be strengthened, tariffs should decrease, and investments should be subsidized. Thus, subsidizing or liberalizing technological trade is a strategic substitute for tougher climate treaties.

Key words: Climate agreements, green technology, dynamic common pool problems, dynamic hold-up problems, incomplete contracts, contract-length
JEL: Q54, Q55, F55, H87, D86

† I am grateful to the audiences at the 2009 AEA meeting, the University of British Columbia, Caltech, CESifo, the Chicago Area Political Economy conference, the University of Copenhagen, Emory University, the Federal Reserve Bank of Chicago, the FEEM conference on Coalitions for Climate Cooperation, Helsinki School of Economics, the Marshall School of Business, the MSRI conference on climate change, Northwestern University, University of Oslo, Princeton University, Paris School of Economics, Simon Fraser University, Stockholm University, and the World Bank. I have benefitted in particular from the comments of Philippe Aghion, Marco Battaglini, David Besanko, Jeff Campbell, Yeon-Koo Che, Rolf Golombek, Faruk Gul, Michael Hoel, Johannes Hörner, Ben Jones, Larry Karp, Charles Kolstad, Matti Liski, Jiang Ning, Wolfgang Pesendorfer, and Roberto Veneziani. Judith N. Levi assisted with the editing.
‡ Northwestern University, University of Oslo, and NBER. Email: harstad@northwestern.edu.

1. Introduction

As demonstrated in Copenhagen, December 2009, implementing efficient climate change policies is going to be tremendously difficult. Not only do independent countries benefit privately from contributing to the public bad, but a lasting solution may require investments in new technology. Recent agreements have two distinct characteristics.[1] First, they have focused on emissions but ignored investments, perhaps because investment levels would be hard to verify by third parties. Second, the commitments are relatively short-term, since committing to the far future may be neither feasible nor desirable. How valuable is such an agreement? How does it affect the incentive to invest in technology, and what characterizes the best agreement?

While these questions are central and important, we do not yet have clear answers, or a good framework for deriving them. This paper attempts to make some progress, and it addresses the questions above head-on by isolating the interaction among negotiations, emissions, and investments. I develop a dynamic framework in which countries pollute as well as invest in technology. The technology reduces the need to pollute, and it can be interpreted as either renewable energy sources or abatement technology. While there is a large number of subgame-perfect equilibria, the Markov perfect equilibria (MPEs) are selected since they are simple and robust. With this refinement, the equilibrium turns out to be unique and the analysis tractable, despite the large number of stocks in the model.

Since the MPE is unique, tacit agreements enforced by trigger strategies are not feasible. But in reality, even domestic stakeholders might act as enforcers if the agreement must be ratified by each country. Although I abstract from domestic politics, I vary the countries' possibilities of negotiating, contracting, and committing, and derive the equilibrium outcome for each situation. Since the equilibrium agreement is also the constrained optimum, the results can be interpreted normatively.

To begin with, countries act noncooperatively at all stages. If one country happens to pollute a lot, the other countries are induced to pollute less in the future since the

[1] The two features characterize the current Kyoto Protocol as well as the recent Copenhagen Accord. The Kyoto Protocol specifies emisson reductions for the five-year period 2008-12, while the Copenhagen Accord lists quantified targets for 2020.

problem is then more severe. They will also invest more in technology to be able to afford the necessary cuts in emissions. If a country invests a lot in abatement technology, it can be expected to pollute less in the future. This induces the other countries to increase their emissions and reduce their own investments. Anticipating these effects, each country pollutes more and invests less than it would in an otherwise similar static model. This dynamic common-pool problem is thus particularly severe.

Short-term agreements on immediate emission levels can nevertheless be worse. A hold-up problem arises when the countries negotiate emissions: if one country has better technology and can cut its emissions fairly cheaply, then its opponents may ask it to bear the lion's share of the burden when collective emissions are reduced.[2] Anticipating this, countries invest less when negotiations are coming up. Consequently, everyone is worse off, particularly if the length of the agreement is short and the number of countries large. This dismal result should provoke us to think hard before recommending a particular climate treaty.

Long-term agreements are better at mitigating the hold-up problem. If commitments are negotiated before a country invests, it cannot be held up by the other countries - at least not before the agreement expires. Thus, countries invest more when the agreement is long-term. Nevertheless, countries underinvest compared to the optimum if the agreement does not last forever. To encourage more investments, the best (and equilibrium) agreement is tougher (in that it stipulates lower emissions) than what is optimal ex post, once the investments are sunk. The equilibrium agreement is derived as a function of its length, while the optimal length is shown to depend on the variance of noncontractible shocks, as well as several other parameters.

The comparative statics are important. In most of the paper, investments are assumed to be noncontractible. This generates the international holdup problem, but it may also lead to domestic holdup problems: Since it is difficult to describe the technology in advance, an innovator will have to develop it first, and then hope that the government is

[2] *Financial Times* reports that "Leaders of countries that want concessions say that nations like Denmark have a built-in advantage because they already depend more heavily on renewable energy" (October 17, 2008: A4). Although the Kyoto Protocol aimed for uniform cuts relative to 1990 levels, exceptions were widespread and there is currently no attempt to harmonize cuts.

willing to pay. If the innovator's intellectual property rights are weak, the government will, in equilibrium, pay less. This "domestic" holdup problem interacts with the international holdup problem, and the optimal agreement reflects them both. If intellectual property rights are weak, investments are low, and a further reduction in investments is particularly harmful. The optimal agreement is then tougher and longer-term, while short-term agreements are more likely to be worse than business as usual.

If technology can be traded or subsidized, then high tariffs or low subsidies discourage investments and, to counteract this, the climate treaty should be tougher and more long-term. The optimal climate treaty is thus a function of trade policies, but the reverse is also true: if the climate treaty is relatively short-term, it is more important to strengthen intellectual property rights, reduce tariffs, and increase subsidies on investments. Negotiating such trade policies is thus a strategic substitute for a tough climate agreement.

By analyzing environmental agreements in a dynamic game permitting incomplete contracts, I contribute to three strands of literature.

The literature on climate policy and environmental agreements is growing.[3] It usually emphasizes the positive effects of regulation on technological change,[4] and a typical recommendation is decade-long short-term agreements, partly to ensure flexibility (see, for example, Karp and Zhao, 2009). The present paper, in contrast, shows that short-term agreements reduce the incentive to invest in new technology and can be worse than business as usual, while long-term agreements are better at mitigating hold-up problems. This builds on Buchholtz and Konrad (1994), who first noted that R&D might decrease prior to negotiations.[5] Beccherle and Tirole (2010) have recently generalized my one-period model and shown that anticipating negotiations can have adverse effects also if the countries, instead of investing, sell permits on the forward market, allow banking, or set

[3] See Kolstad and Toman (2005) on climate policy and Barrett (2005) on environmental agreements. Aldy et al. (2003) and Aldy and Stavins (2007) discuss alternative climate agreement designs.

[4] See, e.g., Jaffe et al. (2003), Newell et al. (2006), Golombek and Hoel (2005). Even when investments are made prior to negotiations, Muuls (2009) finds that investments increase when the negotiations are anticipated. Hoel and de Zeeuw (2009), in contrast, show that R&D can decrease if countries cooperate because they then reduce pollution even without new technology, although there is no negotiation in their model and their analysis hinges on a "breakthrough technology" and binary abatement levels.

[5] Analogously, Gatsios and Karp (1992) show how firms may overinvest prior to merger negotiations.

production standards. With only one period, however, these models miss dynamic effects and thus the consequences for agreement design.

There is already a large literature on the private provision of public goods in dynamic games.[6] Since the evolving stock of public good influences the incentive to contribute, the natural equilibrium concept is Markov perfect equilibrium. As in this paper, equilibrium provision levels tend to be suboptimally low when private provisions are strategic substitutes (Fershtman and Nitzan, 1991; Levhari and Mirman, 1980). There are often multiple MPEs, however, so Dutta and Radner (2009) investigate whether good equilibria, with little pollution, can be sustained by the threat of reverting to a bad one. Since differential games are often hard to analyze, it is quite standard to assume linear-quadratic functional forms,[7] and few authors complicate the model further by adding technological investments. Dutta and Radner (2004) is an interesting exception, but since their costs of pollution and investment are both linear, the equilibrium is "bang-bang" where countries invest either zero or maximally in the first period, and never thereafter. The contribution of this paper is, first, to provide a tractable model, with a unique MPE, in which agents invest as well as pollute over time. This is achieved by assuming that technology has a linear cost and an additive impact. This trick might also be employed when studying industry dynamics, for example, where analytical solutions are rare and numerical simulations typically necessary (see the survey by Doraszelski and Pakes, 2007). Second, incomplete contracts are added to the model. Incomplete contracts are necessary when the question is how agreements on emissions affect the incentive to invest.[8]

By permitting contracts on emissions but not on investments, this paper is in line with the literature on incomplete contracts (e.g., Hart and Moore, 1988). But the standard model has only two stages, and very few papers derive the optimal contract length.

[6] For treatments of differential games, see Başar and Olsder (1999) or Dockner et al. (2000).

[7] For a comprehensive overview, see Engwerda (2005).

[8] To the best of my knowledge, this is the first paper combining incomplete contracts and difference games. Battaglini and Coate (2007) let legislators negotiate spending on transfers and a long-lasting public good. In equilibrium, the legislators contribute too little to the public good, to induce future coalitions to spend more money on it. While the future coalition is unknown, the contract is complete. Hoel (1993) studies a differential game with an emission tax, Yanase (2006) derives the optimal contribution subsidy, Houba et al. (2000) analyze negotiations over (fish) quotas lasting forever, while Sorger (2006) studies one-period agreements. Although Ploeg and de Zeeuw (1992) even allow for R&D, contracts are either absent or complete in all these papers.

Harris and Holmstrom (1987) discuss the length when contracts are costly to rewrite but uncertainty about the future makes it necessary. To preserve the optimal incentives to invest, Guriev and Kvasov (2005) show that the agents should continuously renegotiate the length. Ellman (2006) studies the optimal probability for continuing the contract and finds that it should be larger if specific investments are important. This is somewhat related to my result on the optimal time horizon, but Ellman has only two agents and one investment period, and uncertainty is not revealed over time. In addition, I find that international contracts should be tougher, and more long-term, to compensate for incomplete domestic contracting. Furthermore, the result that short-term agreements can be worse than no agreement is certainly at odds with the traditional literature that focuses on bilateral trade.

Several of the results in this paper survive, qualitatively, in quite general settings. This is confirmed in Harstad (2010), where I also allow for technological spillovers and renegotiation. However, that paper abstracts from uncertainty, intellectual property rights, and trade policies, crucial in the present paper. Finally, by assuming quadratic utility functions, the following analysis goes further and describes when agreements are beneficial and how long they should last.

The model is presented in the next section. Section 3 compares the noncooperative outcome to short-term agreements, and finds conditions under which climate agreements reduce welfare. Section 4 derives the optimal long-term agreement, describing its length and toughness. The effects of subsidies and trade policies are analyzed in Section 5, while Section 6 discusses other extensions and shows that the results survive if, for example, the permits are tradable. After Section 7 concludes, the proofs follow in the Appendix.

2. The Model

2.1. Pollution and Payoffs

Pollution is a public bad. Let G represent the stock of greenhouse gases, and assume that the environmental cost for every country $i \in N \equiv \{1, ..., n\}$ is given by the quadratic cost function:

$$C(G) = \frac{c}{2} G^2.$$

Parameter $c > 0$ measures the importance of climate change.

The stock of greenhouse gases G is measured relative to its natural level. Since the natural level is thus $G = 0$, G tends to approach zero over time (were it not for emissions), and $1 - q_G \in [0, 1]$ measures the fraction of G that "depreciates" every period. G may nevertheless increase if a country's pollution level g_i is positive:

$$G = q_G G_- + \sum_N g_i + \theta. \tag{2.1}$$

By letting G_- represent the stock of greenhouse gases in the previous period, subscripts for periods can be skipped.

The shock θ is arbitrarily distributed with mean 0 and variance σ^2. It has a minor role in the model and most of the results hold without it (i.e., if $\sigma = 0$). However, it is realistic to let the depreciation and cumulation of greenhouse gases be uncertain. Moreover, the main impact of θ is that it affects the future marginal cost of emissions. In fact, the model would be identical if the level of greenhouse gases were simply $\widehat{G} \equiv q_G G_- + \sum_N g_i$ and the uncertainty were instead in the associated cost function. If the cost could be written as $C\left(\widehat{G} + \Theta\right)$, where $\Theta = q_G \Theta_- + \theta$, then $C' = c\Theta + c\widehat{G}$. In either case, a larger θ increases the marginal cost of emissions. Note that, although θ is i.i.d. across periods, it has a long-lasting impact through its effect on G (or on Θ).

The benefit of polluting g_i units is that country i can consume g_i units of energy. Naturally, country i may also be able to consume alternative or renewable energy, depending on its stock of nuclear power, solar technology, and windmills. Let R_i measure this stock and the amount of energy it can produce. The total amount of energy consumed is thus:

$$y_i = g_i + R_i, \tag{2.2}$$

and the associated benefit for i is:

$$B_i(y_i) = -\frac{b}{2}(\overline{y}_i - y_i)^2. \tag{2.3}$$

The benefit function is thus concave and increasing in y_i up to i's bliss point $\overline{y}_i$, which can vary across countries and be a function of time ($\overline{y}_{i,t}$). The bliss point represents the ideal energy level if there were no concern for pollution: a country would never produce more than $\overline{y}_i$ due to the implicit costs of generating, transporting, and consuming energy. The average $\overline{y}_i$ is denoted $\overline{y}$, which also can be time-dependent. Parameter $b > 0$ measures the importance of energy.

Several other interpretations of R_i are consistent with the model. For example, R_i may measure i's abatement technology, i.e., the amount by which i can at no cost reduce (or clean) its potential emissions. If energy production, measured by y_i, is generally polluting, the actual emission level of country i is given by $g_i = y_i - R_i$, implying (2.2), as before. Alternatively, combine (2.2) and (2.3) to let the cost of *abatement* be given by $-b(\overline{y}_i - R_i - g_i)^2/2$. In this case, the marginal cost of abatement is $b(\overline{y}_i - R_i - g_i)$, increasing in i's abatement level ($-g_i$) but decreasing in its abatement technology (R_i).[9]

2.2. Technology and Time

The technology stock R_i may change over time. On the one hand, the technology might depreciate at the expected rate of $1 - q_R \in [0, 1]$. On the other, if r_i measures country i's investment in the current period, then:

$$R_i = q_R R_{i,-} + r_i.$$

As described by Figure 1, the investment stages and the pollution stages alternate over time.[10] Without loss of generality, define "a period" to start with the investment stage and end with the pollution stage. In between, θ is realized. Information is symmetric at all stages.

[9]If, instead, the model focused on technologies that reduced the emission *content* of *each* produced unit (e.g., $g_i = y_i/R_i$), the analysis would be much harder.

[10]This assumption can be endogenized. Suppose the countries can invest at any time in the interval $[t-1, t]$, where t and $t+1$ denote emission stages, but that the investment must take place at least $\xi < 1$ units (measured as a fraction of the period-length) before time t, for the technology to be effective at time t. Then, all countries will invest at time $t - \xi$, never at time $t - 1$.

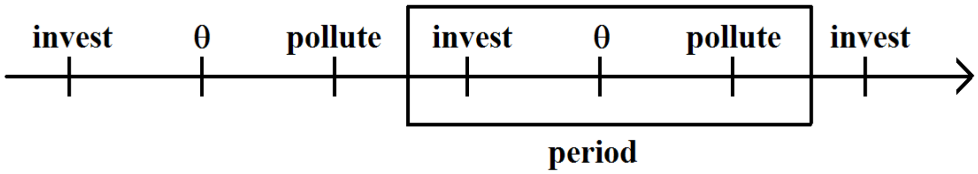

Figure 1: The investment and emission stages alternate over time

Section 5 allows the investments to be verifiable, contractible, and subsidizable. But in most of this paper, the investments are assumed to be observable but not verifiable by third parties. This is in line with the literature on incomplete contracts and may lead to hold-up problems at the international as well as at the national level.

At the international level, it is difficult for countries to negotiate and contract on investment levels. If a country has promised to reduce g_i, it can comply by reducing its short-term consumption or by investing in more long-lasting technology. The difference may be hard to detect by third parties. Therefore, if a country has promised to invest a certain amount, it may be tempted to report other public expenditures as such investments. These problems may explain why the Kyoto Protocol has specified emission quotas, but not investment levels.

At the national level, a government purchasing technology may find it difficult to describe the exact requirements in advance. An innovator or entrepreneur will need to develop the technology first, and then hope the government is willing to pay for it. Let $\mu \in (0, 1]$ be the fraction of the government's benefit that the innovator can capture. Thus, μ represents the innovators' intellectual property right. In particular, μ might be the fraction of an investment that is protectable for the innovator, while the fraction $1 - \mu$ is available for the government to copy for free.[11] But μ may also decrease if the government is a powerful negotiator.[12] If the cost of developing one unit of technology is given by the constant K, then, with free entry, innovators will earn zero profit and charge

[11] This is a slight modification of Acemoglu, Antras, and Helpman (2007). In their model, μ is the fraction of the tasks for which effort can be specified.

[12] Assuming the innovator can set the price, it will capture a fraction μ of the buyer's value. However, if the innovator and the government were bargaining over the price, the innovator would be able to capture the fraction $\widehat{\mu} \equiv \beta\mu$, where $\beta \in (0, 1]$ is the innovator's share of the bargaining surplus (and thus represents its bargaining power). In this case, all the results below continue to hold if only μ is replaced by $\widehat{\mu}$.

the price K.

Consequently, country i's flow utility is:

$$u_i = B_i(y_i) - C(G) - Kr_i,$$

although the equilibrium investment level is given by:

$$\partial \frac{B_i(.) - C(.) + \delta U_{i,+}}{\partial R_i} = \frac{K}{\mu}, \tag{2.4}$$

where $U_{i,+}$ represents country i's continuation value at the start of the next period. The continuation values are given by:

$$U_{i,t} = \sum_{\tau=t}^{\infty} u_{i,\tau}\delta^{\tau-t},$$

where δ is the common discount factor. As mentioned, subscripts denoting period t are often skipped.

This model is more general that might at first appear: For example, if the government is developing technology itself, or if the contract with the developer is complete, then $\mu = 1$, which is a special case of the model, and the results below hold. If the developers of technology are located in foreign countries, the model is unchanged, since such firms earn zero profit in any case. If an innovator can sell (or license) its ideas to several countries at the same time, let K represent the private cost of developing technology that has the potential to raise $\sum_N R_i$ by one unit, and the analysis below needs only small modifications. Tariffs and subsidies on technological trade are discussed in Section 5.

2.3. Definition of an Equilibrium

There is typically a large number of subgame-perfect equilibria in dynamic games, and refinements are necessary. This paper focuses on Markov perfect equilibria (MPEs) where strategies are conditioned only on the pay-off relevant stocks (for further discussion and an exact definition, see Maskin and Tirole, 2001).

There are several reasons for selecting these equilibria. First, experimental evidence suggests that players tend toward Markov perfect strategies rather than supporting the best subgame perfect equilibrium (Battaglini et al., 2010). Second, Markov perfect strategies are simple, since they do not depend on the history in arbitrary ways.[13] This simplifies the analysis as well. Third, focusing on the MPEs is quite standard when studying games with stocks. By doing the same in this paper, its contribution is clarified. Furthermore, in contrast to much of the literature, there is a unique MPE in the present game. This sharpens the predictions and makes institutional comparisons possible. Fifth, the unique MPE coincides with the unique subgame-perfect equilibrium if time were finite but approached infinity. This is particularly important in our context, since the equilibrium is then robust to the introduction of real-world aspects that would make the effective time horizon finite. For example, since fossil fuel is an exhaustible resource, the emission game may indeed have a finite time horizon in the real world. Similarly, politicians' term-limits or short time horizon may force them to view time as expiring.[14] Finally, since the unique MPE makes it impossible to enforce agreements by using trigger strategies, it becomes meaningful to focus instead on settings where countries can negotiate and contract on emission levels - at least for the near future.

I do not explain *why* countries comply with such promises, but one possibility is that the treaty must be ratified domestically and that certain stakeholders have incentives to sue the government unless it complies. By varying the possibilities to negotiate and contract, I derive the outcome for each situation.

At the negotiation stage, I assume the bargaining outcome is efficient and symmetric *if* it should happen that the game and the payoffs are symmetric. This condition is satisfied whether we rely on (i) the Nash Bargaining Solution, with or without side transfers, (ii) the Shapley value, or instead (iii) noncooperative bargaining where one country is randomly selected to make a take-it-or-leave-it offer specifying quotas and transfers. Thus, the condition is quite weak. Note that all countries participate in equilibrium, since there is no stage at which they can close the door to negotiations.

[13] Maskin and Tirole (2001:192-3) defend MPEs since they are "often quite successful in eliminating or reducing a large multiplicity of equilibria," and they "prescribe the simplest form of behavior that is consistent with rationality" while capturing that "bygones are bygones more completely than does the concept of subgame-perfect equilibrium."

[14] More generally, Fudenberg and Tirole (1991:533) suggest that "one might require infinite-horizon MPE to be limits of finite-horizon MPE."

3. Are Agreements Good?

This section discusses the noncooperative outcome as well as the outcome under "short-term" agreements. A comparison reveals that such agreements are not necessarily good.

For future reference, the first-best emission levels are:

$$g_i^*(r) = \overline{y}_i - R_i - \frac{\overline{y}cn^2 + cn(q_G G_- + \theta - R) + \delta K q_G (1 - \delta q_R)}{b + cn^2}, \tag{3.1}$$

which is a function of the stocks in the previous period and - as emphasized - of this period's vector of investments $r \equiv (r_1, ..., r_n)$. Given these emission levels, the first-best investments are:

$$r_i^* = \overline{y} - \frac{q_R R_-}{n} + \frac{q_G G_-}{n} - (1 - \delta q_R)\left(\frac{1 - \delta q_G}{cn^2} + \frac{1}{b}\right) K.$$

Combined,

$$G^* = \sum_N g_i^*(r^*) + q_G G_- = \frac{(1 - \delta q_G)(1 - \delta q_R) K}{cn} + \theta \frac{b}{b + cn^2}. \tag{3.2}$$

Throughout the analysis, I assume $g_i \geq 0$ and $r_i \geq 0$ never bind.[15]

3.1. Business as Usual

Solving the model starts with the following two steps. First, at the beginning of every subgame, one can show that the R_is are payoff-irrelevant, given R: substituting (2.2) into (2.1), we get:

$$G = q_G G_- + \theta + \sum_N y_i - R, \text{ where} \tag{3.3}$$

$$R \equiv \sum_N R_i = q_R R_- + \sum_N r_i. \tag{3.4}$$

This way, the R_is are eliminated from the model: they are *payoff-irrelevant* as long as R is given, and i's Markov perfect strategy for y_i is thus not conditioned on them.[16] A

[15] This is satisfied if $g_i < 0$ and $r_i < 0$ are allowed or, alternatively, if q_G and q_R are small while θ has a limited support. A growing $\overline{y}_t$ is also making it necessary with positive investments. The Appendix (proof of Proposition 1) provides the exact conditions.

[16] That is, there is no reason for one player to condition its strategy on R_i, if the other players are not doing it. Thus, ruling out dependence on R_i is in line with the definition by Maskin and Tirole (2001:202), where Markov strategies are measurable with respect to the coarsest partition of histories consistent with rationality.

country's continuation value U_i is thus a function of G_- and R_-, not of $R_{i,-} - R_{j,-}$; we can therefore write it as $U(G_-, R_-)$, without the subscript i.

Second, the linear investment cost is utilized to prove that the continuation value must be linear in R and, it turns out, in G. Naturally, this simplifies the analysis tremendously.

Proposition 0. (i) *There is a unique symmetric Markov perfect equilibrium.*
(ii) *It is in stationary strategies.*
(iii) *The continuation value is linear in the stocks with the slopes:*

$$\begin{aligned} U_R &= \frac{q_R K}{n}, \\ U_G &= -\frac{q_G K}{n}\left(1 - \delta q_R\right). \end{aligned}$$

This result[17] is referred to as Proposition 0 since it is the foundation for the propositions emphasized below. It holds for all scenarios analyzed in the paper, and for any concave $B(.)$ and convex $C(.)$, even if they are not necessarily quadratic. But to get further explicit results, we impose the quadratic forms.

Proposition 1. *With business as usual, countries pollute too much and invest too little:*

$$r_i^{bau} = \overline{y} - \frac{q_R R}{n} + \frac{q_G G_-}{n} - \frac{k\left(b+cn\right)^2}{cb\left(b+c\right)n} + \frac{\delta U_R\left(b+cn\right)^2}{cb\left(b+c\right)n} - \frac{\delta U_G}{cn} < r_i^*, \quad (3.5)$$

$$g_i^{bau}\left(r^{bau}\right) = \overline{y}_i - R_i - \frac{c\left(n\overline{y} + q_G G_- + \theta - R\right) - \delta U_G}{b+cn} > g_i^*\left(r^{bau}\right) > g_i^*\left(r^*\right). \quad (3.6)$$

The first inequality in (3.6) states that each country pollutes too much compared to the first-best levels, conditional on the investments. A country is not internalizing the cost for the others.

Furthermore, note that country i pollutes less if the existing level of pollution is large and if i possesses good technology, but more if the other countries' technology level is large, since they are then expected to pollute less.

[17] As the proposition states, this is the unique *symmetric* MPE. Since the investment cost is linear, there also exist asymmetric MPEs in which the countries invest different amounts. Asymmetric equilibria may not be reasonable when countries are homogeneous, and they would cease to exist if the investment cost were convex.

In fact, $\overline{y}_i - R_i - g_i^{bau} = \overline{y}_i - y_i$ is the same across countries, in equilibrium, no matter what the differences in technology are. While perhaps surprising at first, the intuition is straightforward. Every country has the same preference (and marginal utility) when it comes to reducing its consumption level relative to its bliss point, and the *marginal* impact on G is also the same for every country: one *more* energy unit generates one unit of emissions. The technology is already utilized to the fullest possible extent, and producing more energy is going to pollute.

Therefore, a larger R, which reduces G, must increase every y_i. This implies that if R_i increases but R_j, $j \neq i$, is constant, then $g_j = y_j - R_j$ must increase. In words: if country i has a better technology, i pollutes less but (because of this) all other countries pollute more. Clearly, this effect reduces the willingness to pay for technology, and generates another reason why investments are suboptimally low, reinforcing the impact of the domestic hold-up problem when $\mu < 1$. The suboptimal investments make it optimal to pollute more, implying the second inequality in (3.6) and a second reason for why pollution is higher than its first-best level.

In sum, a country may want to invest less in order to induce other countries to pollute less and to invest more in the following period. In addition, countries realize that if G_{-} is large for a given R, (3.6) implies that the g_is must decrease. Thus, a country may want to pollute more today to induce others to pollute less (and invest more) in the future. These dynamic considerations make this dynamic common-pool problem more severe than its static counterpart.

3.2. Short-term Agreements

If countries can commit to the immediate but not the distant future, they may negotiate a "short-term" agreement. If the agreement is truly short-term, it is difficult to develop new technology during the time-span of the agreement and the relevant technology is given by earlier installations. This interpretation of short-term agreements can be captured by the timing shown in Figure 2.

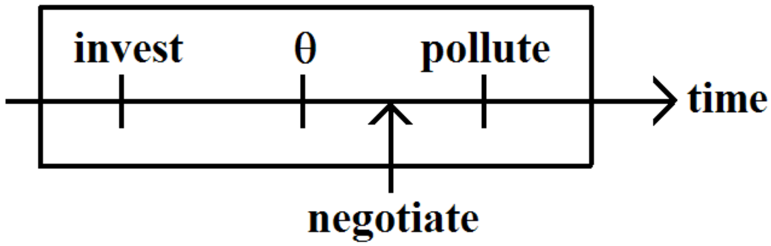

Figure 2: The timing for short-term agreements

Technically, negotiating the g_is is equivalent to negotiating the y_is as long as the R_is are sunk and observable (even if they are not verifiable). Just as in the previous section, (3.3)-(3.4) imply that the R_is are payoff-irrelevant, given R. Even if countries have different R_is, they face the same marginal benefits and costs of reducing y_i relative to $\overline{y}_i$, whether negotiations succeed or not. Symmetry thus implies that y_i is the same for every country in the bargaining outcome. Efficiency requires that the y_is are optimal (all countries agree on this). Consequently, the emission levels are equal to the first-best, conditional on the stocks.

Intuitively, if country i has better technology, its marginal benefit from polluting is smaller, and i is polluting less with business as usual. This gives i a poor bargaining position, and the other countries can offer i a smaller emission quota. At the same time, the other countries negotiate larger quotas for themselves, since the smaller g_i (and the smaller G) reduce the marginal cost of polluting. Anticipating this hold-up problem, every country is discouraged from investing. This international hold-up problem provides a second reason why investments are suboptimally low, in addition to the domestic hold-up problem that arises when $\mu < 1$.

Thus, although emission levels are ex post optimal, actual emissions are larger compared to the first-best levels since the two hold-up problems discourage investments and make it ex post optimal to pollute more.

Proposition 2. *With short-term agreements, countries pollute the optimal amount, given the stocks, but investments are suboptimally low:*

$$\begin{aligned} r_i^{st} &= r_i^* - \left(\frac{n}{\mu} - 1\right)\left(\frac{b+cn^2}{bcn^2}\right)K < r_i^*, \\ g_i^{st}\left(r^{st}\right) &= g_i^*\left(r^{st}\right) > g_i^*\left(r^*\right). \end{aligned}$$

Deriving and describing this outcome is relatively simple because Proposition 0 continues to hold for this case, as proven in the Appendix. In particular, U_G and U_R are exactly the same as in the noncooperative case. This does *not* imply that U itself is identical in the two cases: the levels can be different. But this does imply that when deriving actions and utilities for one period, it is irrelevant whether there will also be a short-term agreement in the next (or any future) period. This makes it convenient to compare short-term agreements to business as usual. For example, such a comparison will be independent of the stocks, since U_G and U_R are identical in the two cases.

3.3. Are Short-Term Agreements Good?

Pollution is less under short-term agreements compared to no agreement. That may not be surprising, since the very motivation for negotiating is to reduce pollution. But what about investments and utilities?

Proposition 3. *Compared to business as usual, short-agreements reduce (i) pollution, (ii) investments, and (iii) utilities if intellectual property rights are weak while the period is short (i.e., if (3.7) holds):*

$$\begin{aligned} \mathrm{E}g^{st}\left(r^{st}\right) &= \mathrm{E}g^{bau}\left(r^{bau}\right) - K\left(\frac{1}{\mu} - \frac{\delta q_R}{n}\right)\frac{n-1}{n\,(b+c)}, \\ r_i^{st} &= r_i^{bau} - K\left(\frac{1}{\mu} - \frac{\delta q_R}{n}\right)\frac{(n-1)^2}{n\,(b+c)}, \end{aligned}$$

$$U^{st} < U^{bau} \Leftrightarrow \left(\frac{n}{\mu} - 1\right)^2 - (1-\delta q_R)^2 > \sigma^2 \frac{(b+c)\,(bcn/K)^2}{(b+cn^2)\,(b+cn)^2}. \tag{3.7}$$

Rather than being encouraging, short-term agreements impair the motivation to invest. The reason is the following. In anticipating of negotiations, the hold-up problem is exactly as strong as the crowding-out problem in the noncooperative equilibrium; in either case, each country enjoys only $1/n$ of the total benefits generated by its investments. In addition, when an agreement is expected, i understands that pollution will be reduced. A further decline in emissions, made possible by new technology, is then less valuable. Hence, each country is willing to pay less for technology.[18]

Since investments decrease under short-term agreements, utilities can decrease as well. This is the case, in particular, if investments are important because they are already well below the optimal level. Thus, short-term agreements are bad if intellectual property rights are weak (μ small), the number of countries is large, and the period for which the agreement lasts is very short. If the period is short, δ and q_R are large, while the uncertainty from one period to the next, determined by σ, is likely to be small. All changes make (3.7) reasonable, and it always holds when the period is very short ($\sigma \to 0$).

But at the emission stage, once the investments are sunk, all countries benefit from negotiating an agreement. It is the *anticipation* of negotiations which reduces investments and perhaps utility. Thus, if (3.7) holds, the countries would have been better off if they could commit to not negotiating short-term agreements. In particular, it may be better to commit to emission levels before the investments occur.

4. The Optimal Agreement

The hold-up problem under short-term agreements arises because the g_is are negotiated after investments are made. If the time horizon of an agreement is longer, however, it is possible for countries to develop technologies within the time frame of the agreement. The other countries are then unable to hold up the investing country, since the quotas have already been agreed to, at least for the near future.

To analyze such long-term agreements, let the countries negotiate and commit to emission quotas for T periods. The next subsection studies equilibrium investment, as a function of such an agreement. Taking this function into account, the second subsection

[18] A counter-argument is that, if an agreement is expected, it becomes more important to invest to ensure a decent energy consumption level. This force turns out to be smaller, at least for quadratic utility functions.

derives the optimal (and equilibrium) emission quotas, given T. Finally, the optimal T is characterized.

If the agreement is negotiated just before the emission stage in a period, then the quotas and investments for that period are given by Proposition 2, above: the quotas will be first-best, given the stocks, but investments are too low, due to the hold-up problems.

For the subsequent periods, it is irrelevant whether the quotas are negotiated before the first emission stage, or instead at the start of the next period, since no information is revealed, and no strategic decisions are made, in between. To avoid repeating earlier results, I will focus on the subsequent periods, and thus implicitly assume that the T-period agreement is negotiated at the start of period 1, as described by Figure 3.

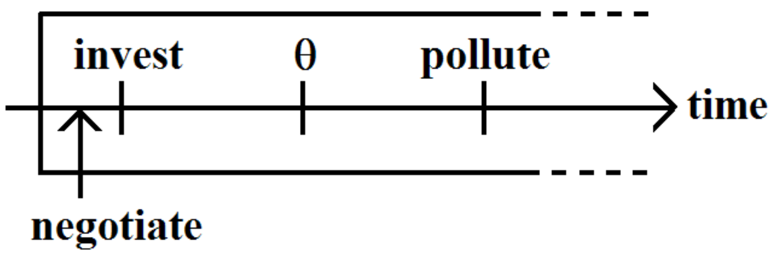

Figure 3: The timing for long-term agreements

4.1. Investments as a Function of the Quotas

When investing in period $t \in \{1, 2, ..., T\}$, countries take the quotas, the $g_{i,t}^{lt}$s, as given, and the continuation value for period $T+1$ is $U(G_T, R_T)$. A country is willing to pay more for innovations and investments if its quota, $g_{i,t}$, is small, since it is going to be very costly to comply if the sum $y_{i,t} = g_{i,t} + R_{i,t}$ is also small. Anticipating this, innovations and investments decrease in $g_{i,t}$.

However, compared to the investments that are first-best conditional on the quotas, $r_{i,t}^{*}(g_{i,t})$, equilibrium investments are likely to be too low. In every period, the innovators fear to be held up, if $\mu < 1$, and thus they invest and innovate only up to the point where the countries' willingness to pay for $1 - \mu$ units equals the cost of developing one unit of technology. Furthermore, a country anticipates that having good technology will worsen its bargaining position in the future, once a new agreement is to be negotiated. At that stage, having good technology leads to a lower $g_{i,t}$ since the other countries can hold up country i when it is cheap for i to reduce its emissions.[19] Anticipating this, countries invest less in the last period, particularly if that period is short (δ large), the technology long-lasting (q_R large), and the number of countries large (n large).

PROPOSITION 4. (i) *Investments increase in μ but decrease in $g_{i,t}$.*
(ii) *Investments are suboptimally low if $\mu < 1$, for any given quota and period.*
(iii) *In the last period, investments are suboptimally low if $\mu < 1$ or $\delta q_R > 0$:*

$$\begin{aligned}
r_{i,t}^{*}(g_{i,t}) &= \overline{y}_i - q_R R_{i,-} - g_{i,t} - (1-\delta q_R) K/b \\
&\geq \text{ (strict if } \mu < 1) \\
r_{i,t}(g_{i,t}) &= r_{i,t}^{*}(g_{i,t}) - \left(\frac{1}{\mu} - 1\right)\frac{K}{b} \text{ for } t < T \\
&\geq \text{ (strict if } \delta q_R > 0) \\
r_{i,t}(g_{i,t}) &= r_{i,t}^{*}(g_{i,t}) - \left(\frac{1}{\mu} - 1 + \delta q_R \left(1 - \frac{1}{n}\right)\right)\frac{K}{b} \text{ for } t = T.
\end{aligned}$$

4.2. The Optimal Emission Quotas

At the emission stage, the (ex post) optimal pollution level is given by $g_i^*\left(r^{lt}\right)$, as before. However, the countries anticipate that the negotiated $g_{i,t}$s are going to influence investments in technology: the smaller the quotas, the larger the investments. Thus, since the investments are suboptimally low, the countries have an incentive to commit to quotas that are actually smaller than the expectation of $g_i^*\left(r^{lt}\right)$, just to encourage investments. The smaller equilibrium investments are compared to the optimal investments, the lower are the negotiated $g_{i,t}^{lt}$s, compared to emission levels that are ex post optimal.

PROPOSITION 5. (i) *If $\mu < 1$, the negotiated quotas are strictly smaller than the ex post optimal levels, in every period.*
(ii) *For the last period, the negotiated quotas are strictly smaller than the ex post optimal quotas if either $\mu < 1$ or $\delta q_R > 0$.*

[19] Or, if no agreement is expected in the future, a large $R_{i,T+1}$ reduces $g_{i,T+1}$ and increases $g_{j,T+1}$, as proven in Section 3.1.

(iii) *The negotiated emission levels are identical to the emission levels that would have been first-best if investments had been first-best*:

$$g_{i,T}^{lt} = \mathrm{E}g_{i,T}^{*}(r^{*}) = \mathrm{E}g_{i,T}^{*}\left(r^{lt}\right) - \frac{1/\mu - 1 + \delta q_R (1 - 1/n)}{b + cn^2} K \text{ for } t = T, \quad (4.1)$$

$$g_{i,t}^{lt} = \mathrm{E}g_{i,t}^{*}(r^{*}) = \mathrm{E}g_{i,T}^{*}\left(r^{lt}\right) - \frac{1/\mu - 1}{b + cn^2} K \text{ for } t < T. \quad (4.2)$$

Parts (i)-(ii) have the following intuition. If μ is small, the last terms of (4.1)-(4.2) are large, and every $g_{i,t}^{lt}$ must decline relative to $g_i^*\left(r^{lt}\right)$. This makes the agreement more demanding or *tougher* to satisfy at the emission stage. The purpose of such a seemingly overambitious agreement is to encourage investments, since these are suboptimally low when μ is small. Encouraging investments is especially important in the last period, since investments are particularly low then, according to Proposition 4. Thus, the optimal agreement is tougher to satisfy over time.[20]

On the other hand, if $\mu = \delta q_R = 0$, the last terms of (4.1)-(4.2) are zero, meaning that the commitments under the best long-term agreement also maximize the expected utility ex post. In this case, there are no underinvestments, and there is no need to distort $g_{i,t}^{lt}$ downwards.

Part (iii) can be explained as follows. While the attempt to mitigate underinvestments reduces $g_{i,t}^{lt}$ compared to $g_i^*\left(r^{lt}\right)$, the fact that investments are low implies that it is ex post optimal to pollute more, so $g_i^*\left(r^{lt}\right)$ *increases* relative to $g_i^*(r^*)$. These two effects turn out to cancel.[21]

Just as in the previous cases, it turns out that the continuation value U is linear in the stocks, making the analysis tractable. Moreover, $U_R^{lt} = q_R K/n$ and $U_G^{lt} = -q_G(1 - \delta q_R) K/n$, as before. The predicted contract and investments are therefore robust to whether there is a long-term agreement, a short-term agreement, or no agreement in the subsequent period.

[20] This conclusion would be strengthened if the quotas were negotiated just before the emission stage in the first period. Then, the first-period quotas would be ex post optimal since these quotas would, in any case, have no impact on investments. It is easy to show that these quotas are expected to be larger than the quotas described by Proposition 5 - whether or not this is conditioned on investment levels.

[21] The reason is that, in this equilibrium as well as in the first-best outcome, y_i is independent of g_i, so a smaller g_i is only reducing G and increasing R_i. Since the marginal cost of increasing R_i is constant, the optimal G is the same in this equilibrium and in the first-best outcome.

4.3. The Optimal Agreement Length

The length of the agreement might be limited by the countries' ability to commit. Otherwise, the optimal and equilibrium T trades off two concerns. On the one hand, investments are particularly low at the end of the agreement, before a new agreement is to be negotiated. This hold-up problem arises less frequently, and is delayed, if T is large. On the other hand, the stochastic shocks cumulate over time, and they affect the future marginal costs of pollution. This makes it hard to estimate the optimal quotas for the future, particularly when T is large.

In general, the optimal length of an agreement depends on the regime that is expected to replace it. This is in contrast to the other contracts studied above, which have been independent of the future regime. When the time horizon is chosen, it is better to commit to a longer-term agreement if everyone expects that, once it expires, the new regime is going to be worse (e.g., business as usual).

On the other hand, if future as well as present negotiators are able to commit to future emissions, then we can anticipate that the next agreement is also going to be optimal. Under this assumption, the optimal agreement is derived and characterized in the Appendix.

PROPOSITION 6. (i) *The agreement's optimal length T^* decreases in μ, b, c, and σ, but increases in n, q_R, and K.*
(ii) *In fact, $T^* = \infty$ if:*

$$q_R K^2 \frac{1-1/n}{b}\left[\frac{1}{\mu} + 3\delta q_R \frac{1-1/n}{2}\right] \geq \frac{c\sigma^2 q_G^2}{2\left(1-q_G^2\right)\left(1-\delta q_G^2\right)}.$$

If θ were known or contractible, the agreement should last forever. Otherwise, the length of the agreement should be shorter if future marginal costs are uncertain (σ large) and important (c large). However, a larger T is preferable if the underinvestment problem is severe. This is the case if the intellectual property rights are weak (μ small), the technology is long-lasting (q_R large), and the number of countries is large. If b is large while K is small, then consuming energy is much more important than the concern for future bargaining power. The hold-up problem is then small, and the optimal T declines.

5. R&D Policies and Climate Agreements

So far, investments in technology have been noncontractible. But since, as a consequence, investments were suboptimally low, the countries have incentives to search for ways by which investments can be subsidized. This section allows for such subsidies and shows that the framework continues to provide important lessons.

Let $\phi \in [0,1)$ be an ad valorem subsidy captured by the innovator or developer of technology. It may denote the share of research expenses borne by the government (as in Grossman and Helpman, 1991:264). As before, K is the cost of increasing R_i by one unit, while $\mu \in (0,1]$ is the fraction of the purchaser's benefit that can be captured by the seller. With free entry of innovators, the equilibrium investments will be given by the following condition (replacing (2.4)):

$$\partial \frac{B_i\left(.\right) - C\left(.\right) + \delta U_{i,+}}{\partial R_i} = \frac{K\left(1-\phi\right)}{\mu}. \tag{5.1}$$

If the typical developer of technology for one country is located abroad, then ϕ can be interpreted as an import subsidy. If $\phi < 0$, then $\tau \equiv -\phi > 0$ may be interpreted as an import tariff.[22]

All the proofs in the Appendix are derived as a function of ϕ. As suggested by (5.1), the effect of ϕ is similar to the effect of μ. If the subsidy is exogenous and low, or the tariff $\tau \equiv -\phi$ is high, then investments decline. A further reduction in investment is then particularly bad, which implies that short-term agreements are worse than business as usual. To encourage more investments, the best climate agreement is tougher and longer-term.

Proposition 7. *If the subsidy ϕ is low or the tariff $\tau \equiv -\phi$ is high, the optimal*

[22] Technically, this requires the tariff, or the import subsidy, to be proportional to the cost of developing technology. If, instead, the import subsidy ϕ', or the tariff $-\phi'$, were proportional to the sales value, (5.1) should be:

$$\partial \frac{B_i\left(.\right) - C\left(.\right) + \delta U_{i,+}}{\partial R_i} = \frac{K}{\mu\left(1+\phi'\right)}.$$

Alternatively, with an estate subsidy ϕ'', or tariff $-\phi''$, (5.1) should be:

$$\partial \frac{B_i\left(.\right) - C\left(.\right) + \delta U_{i,+}}{\partial R_i} = \frac{K}{\mu} - \phi''.$$

In these cases, the effects of ϕ' and ϕ'' would be similar to the effects of ϕ.

agreement is tougher and more long-term, while short-term agreements are likely to be worse than business as usual.

If ϕ, τ, or μ can be specified by international law, one may ask for their ideal levels. In particular, how do the optimal subsidy, tariff, and intellectual property right protection depend on the climate treaty?

PROPOSITION 8. *The optimal ϕ and μ are larger if the agreement is short-term or absent. They are given by:*
(i) *Equation (5.2) for short-term agreements as well as for business as usual;*
(ii) *Equation (5.3) for a long-term agreement's last period;*
(iii) *Equation (5.4) for a long-term agreement, except for its last period:*

$$\phi^*_{st} = \phi^*_{bau} = 1 - \mu/n > \tag{5.2}$$

$$\phi^*_{lt,T} = 1 - \mu\left[1 - \delta q_R\left(1 - 1/n\right)\right] > \tag{5.3}$$

$$\phi^*_{lt,t} = 1 - \mu. \tag{5.4}$$

If the climate treaty is short-term, the hold-up problem is larger and it is more important to encourage investments by protecting intellectual property rights, subsidizing technological trade, and reducing tariffs. Such trade agreements are thus strategic substitutes for climate treaties: weakening cooperation in one area makes further cooperation in the other more important. As before, the optimal agreement is also going to be the equilibrium when the countries negotiate, since they are symmetric at the negotiation stage w.r.t. $\overline{y}_{i,t} - y_{i,t}$, no matter what their technological differences are.

As suggested by Proposition 7, the optimal agreement has a shorter length if the subsidy is large. If the subsidy can be freely chosen and set in line with Proposition 8, short-term agreements are actually first-best: while the optimal subsidy induces first-best investments, the negotiated emission levels are also first-best, conditional on the investments. Long-term agreements are never first-best, however, due to the stochastic θ.

PROPOSITION 9. *If ϕ and μ can be set according to Proposition 8, short-term agreements implement the first-best outcome, but long-term agreements do not.*

6. Robustness and Limitations

This paper has focused on the interaction between investments in technology and climate agreements on emissions. To isolate these effects, the model abstracted from a range of real-world complications. While some assumptions have been crucial for the results, others can easily be relaxed.

For example, the discussion has ignored trade in emission allowances, presuming that such trade is prohibited. This assumption is not necessary, however.

PROPOSITION 10. *Suppose the emission allowances are tradable.*
(i) *All results survive, whether or not side payments are available at the negotiation stage.*
(ii) *In every period, the equilibrium permit price is $b(\overline{y}_i - y_{i,t})$, which decreases in μ but increases toward the end of the agreement.*

The permit price is $b(\overline{y}_i - y_{i,t}) = b(\overline{y}_i - g_{i,t} - R_{i,t})$, i.e., the marginal benefit of polluting one more unit, keeping the stocks constant. The larger $R_{i,t}$ is, the lower the marginal benefit of polluting is, and thus the permit price. It follows that the equilibrium permit price is larger if intellectual property rights are weak, as well as in the last period. Under a short-term agreement, which is negotiated taking technologies as given, the permit price declines in the allocated quotas. This is not true for long-term agreements, however, since a smaller $g_{i,t}^{lt}$ will raise $R_{i,t}$ by the very same amount: $\partial r_{i,t}^{lt}/\partial g_{i,t}^{lt} = -1$.

The quadratic functional forms are necessary when comparing utilities (as in Proposition 3) and when deriving the optimal length of an agreement (as in Proposition 6). Many of the other results, however, continue to hold, qualitatively, for any concave $B_i(.)$ and convex $C(.)$, even if they are not quadratic. This is confirmed in Harstad (2010), which also allows for technological spillovers. In addition, that paper permits the commitments to be renegotiated over time, and shows that this might be beneficial.[23]

The paper has allowed for certain types of heterogeneity. In particular, countries can have different initial technology, and their bliss point for energy consumption may vary.

[23] Renegotiation is nevertheless ignored in the present paper, since the possibility to renege may undermine the ability to commit in the first place. While Harstad (2010) assumes that the threat point under renegotiating is the existing agreement, this requires some discipline. If negotiators can credible threaten to exit the agreement if renegotiation fails, then allowing for renegotiation is equivalent to a sequence of short-term agreements, studied above.

Other types of heterogeneity would be harder to analyze. For example, suppose the cost of developing technology, K, varied across countries. In equilibrium, only countries with a small K would invest. This would also be optimal, but, just as before, the investing countries would invest too little. In a long-term agreement, one could encourage these countries to invest more by reducing g_i^{lt}. Such small g_is would not be necessary (or optimal) for noninvesting countries. Naturally, the investing countries may require some compensation to accept the small quotas. At the same time, a small g_i would not motivate country i to invest if i were allowed to purchase permits from noninvesting countries with more permits. Thus, with heterogeneous investment costs, Proposition 10 would be false. Evaluating political instruments under heterogeneity is thus an important task for future research.

The model above has predicted full participation in a climate treaty. This followed since there was no stage at which countries could opt out of the negotiation process. If such a stage were added to the model, free-riding may emerge. For example, in the one-period model analyzed by Barrett (2005), only three countries participate in equilibrium, when utility functions are quadratic. One may conjecture, however, that the number could be larger in a dynamic model, like the one above. If just a few countries decided to participate, they may find it optimal to negotiate short-term agreements, rather than long-term agreements, in the hope that the nonparticipants will join later. Since the participants invest less under short-term agreements, this credible threat might discourage countries when considering to free-ride.

7. Conclusions

While mitigating climate change will require emission reduction as well as the development of new technology, recent agreements have focused on short-term emissions. What is the value of such an agreement? How does it influence the incentive to invest, and what is the best agreement?

To address these questions, this paper provides a framework where countries over time both pollute and invest in environmentally friendly technologies. The analysis generates a number of important lessons.

First, the noncooperative outcome is particularly bad. With business as usual, countries pollute too much, not only because they fail to internalize the externality, but also because polluting now motivates the other countries to pollute less and invest more in the future. Similarly, countries invest too little in technology, to induce the others to invest more and pollute less themselves.

Second, short-term agreements can, nevertheless, be worse. At the negotiation stage, a country with good technology is going to be held up by the others, requiring it to reduce its pollution a great deal. Anticipating this, countries invest less when negotiations are coming up. This makes the countries worse off relative to business as usual, particularly if the agreement has a short duration and intellectual property rights are poorly enforced.

Third, the optimal agreement described. A tough agreement, if long-term, encourages investments. Thus, if the countries can commit to the far future, the optimal and equilibrium agreement is tougher and more longer-term if, for example, technologies are long-lasting but intellectual property rights weak.

Fourth, the optimal climate treaty is a function of trade policies. If technologies can be traded or subsidized, high tariffs and low subsidies discourage investments and, to counteract this, the climate treaty should be tougher and more long-term. If the climate treaty is absent or relatively short-lasting for exogenous reasons, then tariffs should decrease, intellectual property rights should be strengthened, and investments or trade in technology should be subsidized. Negotiating such policies is thus a strategic substitute to a tough climate treaty: if one fails, the other is more important.

The benchmark model is intentionally simple. In isolating the interaction between negotiations, emissions, and technological investments, it illuminates the challenges that would arise *even if* countries were similar, information complete, participation full, and the countries capable of committing to their promises. Future research should relax these assumptions to deepen our understanding of good climate policies.

8. Appendix

The following proofs allow for the subsidy ϕ, introduced in Section 5 (Propositions 0-6 follows if $\phi = 0$). While U_i is the continuation value for a subgame starting with the investment stage, let W_i represent the (interrim) continuation value at (or just before) the emission stage. To shorten equations, use $m \equiv -\delta \partial U_i / \partial G_-$, $z \equiv \delta \partial U_i / \partial R_-$, $\widetilde{R} \equiv q_R R_-$, $\widetilde{G} \equiv q_G G_- + \theta$ and $\widetilde{y}_i \equiv y_i + \overline{y} - \overline{y}_i$, where $\overline{y} \equiv \sum_N \overline{y}_i / n$. The proof for the first-best is omitted since it would follow the same lines as the following proofs.

Proof of Proposition 0.
Just before the emission stage, θ is known and the payoff-relevant states are R and $\widetilde{G}$.[24] A country's (interrim) continuation value is $W(\widetilde{G}, R)$. Anticipating this, equilibrium investments are given by:

$$\frac{\partial \mathrm{E}W(\widetilde{G}, \sum_N R_i)}{\partial R_i} = \frac{\partial \mathrm{E}W(\widetilde{G}, \sum_N R_i)}{\partial R} = k \equiv \frac{K(1-\phi)}{\mu}, \tag{8.1}$$

where expectations are taken w.r.t. θ. This implies, since the marginal cost of increasing R is constant, that the equilibrium R must be independent of R_-. Thus, when all countries invest the same, a marginally larger R_- implies that R will be unchanged, but r_i can decline by q_R/n units. It follows that:

$$\frac{\partial U}{\partial R_-} = \frac{q_R K}{n}. \tag{8.2}$$

At the emission stage, a country's first-order condition for y_i is:

$$0 = B'(\widetilde{y}_i - \overline{y}) - C'\left(\widetilde{G} - R + \sum_N \widetilde{y}_j\right) + \delta U_G(\widetilde{G} - R + \sum_N \widetilde{y}_j, R), \tag{8.3}$$

implying that all $\widetilde{y}_i$s are identical. From (8.2), we know that $U_{RG} = U_{GR} = 0$, and U_G cannot be a function of R. Therefore, (8.3) implies that $\widetilde{y}_i$, G and thus $B(\widetilde{y}_i - \overline{y}) - C(G) \equiv \gamma(.)$ are functions of $\widetilde{G} - R$ only. Hence, write $G\left(\widetilde{G} - R\right)$. Then, (8.1) becomes:

$$\frac{\partial \mathrm{E}\left[\gamma\left(q_G G_- + \theta - R\right) + \delta U\left(G\left(q_G G_- + \theta - R\right), R\right)\right]}{\partial R} = k.$$

This requires $q_G G_- - R$ to be a constant, say ξ, which is independent of the stocks. This implies that $\partial r_i / \partial G_- = q_G / n$ and U becomes:

$$\begin{aligned}
U(G_-, R_-) &= \mathrm{E}\gamma(\xi + \theta) - Kr + \mathrm{E}\delta U\left(G(\xi + \theta), R\right) \\
&= \mathrm{E}\gamma(\xi + \theta) - K\left(\frac{q_G G_- - \xi - q_R R_-}{n}\right) + \mathrm{E}\delta U\left(G(\xi + \theta), q_G G_- - \xi\right) \Rightarrow \\
\partial U / \partial G_- &= -K\left(\frac{q_G}{n}\right) - \delta U_R q_G = -\frac{K q_G}{n}(1 - \delta q_R).
\end{aligned}$$

[24] As explained in the text, there is no reason for one country, or one firm, to condition its strategy on R_i, given R, if the other players are not doing it. Ruling out such dependence is consistent with the definition of Markov and Tirole (2001).

Proof of Proposition 1.
From (8.3),

$$\widetilde{y}_i = \overline{y} - \frac{m + cG}{b} \Rightarrow y_i = \overline{y}_i - \frac{m + cG}{b} \Rightarrow \tag{8.4}$$

$$G = \sum_N (y_i - R_i) + \widetilde{G} = \widetilde{G} - R + n\left(\overline{y} - \frac{m + cG}{b}\right) = \frac{b\overline{y}n - mn + b\left(\widetilde{G} - R\right)}{b + cn} \tag{8.5}$$

$$y_i = \overline{y}_i - \frac{m}{b} - \frac{c}{b}\left(\frac{b\overline{y}n - mn + b\left(\widetilde{G} - R\right)}{b + cn}\right) = \overline{y}_i - \frac{c\overline{y}n + c\left(\widetilde{G} - R\right) + m}{b + cn} \Rightarrow$$

$$g_i = y_i - R_i = \overline{y}_i - \frac{c\overline{y}n + c\left(\widetilde{G} - R\right) + m}{b + cn} - R_i,$$

A comparison to the first-best gives (3.6). Interrim utility (after investments are sunk) can be written as:

$$W_i^{no} \equiv -\frac{c}{2}G^2 - \frac{b}{2}\left(\overline{y}_i - y_i\right) + \delta U(G, R) = -\frac{c}{2}\left(1 + \frac{c}{b}\right)G^2 - \frac{Gmc}{b} - \frac{m^2}{2b} + \delta U(G, R).$$

Since $\partial G/\partial R = -b/\left(b + cn\right)$ from (8.5), equilibrium investments are given by:

$$\mathrm{E}\partial W_i^{no}/\partial R = c\left(1 + \frac{c}{b}\right)\left(\frac{b}{b + cn}\right)\mathrm{E}G + \frac{bm\left(1 + c/b\right)}{b + cn} + z = k. \tag{8.6}$$

Taking expections of G in (8.5), subsituting in (8.6) and solving for R gives:

$$R = \overline{y}n + \mathrm{E}\widetilde{G} - k\frac{\left(b + cn\right)^2}{bc\left(b + c\right)} + \frac{m}{c} + z\frac{\left(b + cn\right)^2}{bc\left(b + c\right)} \Rightarrow \tag{8.7}$$

$$r_i = \frac{R - q_R R_-}{n} = \overline{y} + \frac{q_G G_-}{n} - k\frac{\left(b + cn\right)^2}{bc\left(b + c\right)n} + \frac{m}{c} + z\frac{\left(b + cn\right)^2}{bc\left(b + c\right)n}.$$

A comparison to the first-best gives (3.5).

In steady state,

$$\begin{aligned} G &= \frac{b\overline{y}n - mn}{b + cn} + \frac{b}{b + cn}\left(\theta + k\frac{\left(b + cn\right)^2}{bc\left(b + c\right)} - \frac{m}{c} - z\frac{\left(b + cn\right)^2}{bc\left(b + c\right)} - \overline{y}n\right) \\ &= \frac{b}{b + cn}\theta + \frac{\left(b + cn\right)}{c\left(b + c\right)}\left(k - z\right) - \frac{m}{c}, \\ R &= \overline{y}n - \frac{\left(b + cn\right)^2}{bc\left(b + c\right)}\left(k - z\right) + \frac{m}{c} + q_G\left[\frac{b}{b + cn}\theta_- + \frac{\left(b + cn\right)}{c\left(b + c\right)}\left(k - z\right) - \frac{m}{c}\right] \\ &= \overline{y}n - \left(k - z\right)\frac{\left(b + cn\right)}{c\left(b + c\right)}\left[\frac{b + cn}{b} - q_G\right] + \frac{m}{c}\left(1 - q_G\right) + \frac{q_G b\theta_-}{b + cn}. \end{aligned}$$

If the support of θ is $\left[\underline{\theta}, \overline{\theta}\right]$, investment levels are *always* positive if:

$$\begin{aligned}0 \quad &\leq \quad \max_{\theta_t, \theta_{t-1}} R - q_R R_- = (1-q_R)\left[-\left(k - \frac{\delta q_R K}{n}\right)\frac{(b+cn)}{c(b+c)}\left[\frac{b+cn}{b} - q_G\right] + \frac{\delta q_G(1-\delta q_R)}{c}(1-q_G)\right] \\ &+\frac{q_G b}{b+cn}\left(\underline{\theta} - q_R\overline{\theta}\right) + n\overline{y}_t - q_R n\overline{y}_{t-1},\end{aligned}$$

while the emission level is *always* positive if:

$$0 \leq \min_{\theta_t, \theta_{t-1}} G - q_G G_- - \theta = (1-q_G)\left[\frac{(b+cn)}{c(b+c)}(k-z) - \frac{m}{c}\right] - \frac{cn + q_G b}{b+cn}\overline{\theta}.$$

Proof of Proposition 2.
At the emission stage, the countries negotiate the g_is. g_i determines $\widetilde{y}_i$, and since countries have symmetric preferences over $\widetilde{y}_i$ (in the negotiations as well as in the default outcome), the $\widetilde{y}_i$s must be identical in the bargaining outcome and efficiency requires:

$$0 = B'\left(\widetilde{y}_i - \overline{y}\right)/n - C'\left(\widetilde{G} - R + \sum \widetilde{y}_i\right) + \delta U_G(\widetilde{G} - R + \sum \widetilde{y}_i, R). \qquad (8.8)$$

The rest of the proof of Proposition 0 continues to hold: R will be a function of G_- only, so $U_{R_-} = q_R K/n$. This makes $\mathrm{E}\widetilde{G} - R$ a constant and $U_{G_-} = -q_G(1-\delta q_R)K/n$, just as before. The comparative static becomes the same, but the *levels* of g_i, y_i, r_i, u_i and U_i are obviously different from the previous case.

The first-order condition (8.8) becomes:

$$\begin{aligned}0 &= -ncG + b\overline{y} - b\widetilde{y}_i - nm \Rightarrow y_i = \overline{y}_i - \frac{nm + ncG}{b}. \\ G &= \widetilde{G} + \sum_j (y_j - R_j) = \widetilde{G} + n\left(\overline{y} - \frac{nm+ncG}{b}\right) - R \Rightarrow \\ G &= \frac{b\overline{y}n - mn^2 + b\left(\widetilde{G} - R\right)}{b + cn^2}. \qquad (8.9)\end{aligned}$$

Interrim utility is

$$W_i^{st} = -\frac{c}{2}G^2 - \frac{b}{2}\left(\frac{nm+ncG}{b}\right)^2 + \delta U(G,R).$$

Since (8.9) implies $\partial G/\partial R = -b/(b+cn^2)$, equilibrium investments are given by:

$$\begin{aligned}k &= \mathrm{E}\frac{\partial W_i^{st}}{\partial R} = \mathrm{E}G\left(c + \frac{c^2n^2}{b}\right)\left(\frac{b}{b+cn^2}\right) + \frac{cmn^2}{b}\left(\frac{b}{b+cn^2}\right) + m\left(\frac{b}{b+cn^2}\right) + z \\ &= c\mathrm{E}G + m + z.\end{aligned}$$

Subsituted in (8.9), after taking the expection of it, and solving for R, gives

$$R^{st} = q_G G_- + n\overline{y} + \frac{m}{c} - \left(\frac{b+cn^2}{b}\right)\left(\frac{k}{c} - \frac{z}{c}\right). \qquad (8.10)$$

The proof is completed by comparing r_i^* to $r_i^{st} = (R^{st} - q_R R_-)/n$, which is:

$$
\begin{aligned}
r_i^{st} &= \overline{y} - \frac{q_R R_-}{n} + \frac{q_G G_-}{n} - \left(\frac{b+cn^2}{bcn}\right)(k - \delta U_R) - \frac{\delta U_G}{cn} \\
&= r_i^* - K\left(\frac{b+cn^2}{bcn^2}\right)\left(\frac{nk}{K} - 1\right) < r_i^*.
\end{aligned}
$$

Proof of Proposition 3.
Part (i) and (ii) follow after some algebra when comparing emissions and investments for business as usual to short-term agreements. Substituted in u_i, which in turn should be substituted in $U = u_i + \delta U_+(.)$, allows us to compare U^{bau} and U^{st}. Then, straightforward algebra gives part (iii).

With business as usual, since $\widetilde{G} = q_G G_- + \theta$, (8.5) gives $G = \mathrm{E}G + \theta b/(b+cn)$. Substituted in (8.4) gives:

$$
y_i = \overline{y}_i - \frac{m + cG}{b} = \overline{y}_i - \frac{(k-z)(b+cn)}{b(b+c)} - \frac{\theta c}{b+cn}.
$$

This is helpful when calculating u_i^{bau}. It becomes:

$$
\begin{aligned}
u_i^{bau} &= -\frac{c}{2}\left(\frac{k(b+cn)}{c(b+c)} - \frac{m}{c} - \frac{z(b+cn)}{c(b+c)} + \frac{\theta b}{b+cn}\right)^2 - \frac{b}{2}\left(\frac{(k-z)(b+cn)}{b(b+c)} + \frac{\theta bc}{b(b+cn)}\right)^2 \\
&\quad -\frac{K}{n}\left(-\widetilde{R} + q_G G_- - \frac{k(b+cn)^2}{cb(b+c)} + \overline{y}n + \frac{z(b+cn)^2}{cb(b+c)} + \frac{m}{c}\right) \Rightarrow \\
\mathrm{E}u_i^{bau} &= -\frac{1}{2}(k-z)^2\left(\frac{b+cn}{b+c}\right)^2\left(\frac{1}{c} + \frac{1}{b}\right) - \frac{m^2}{2c} + \frac{m}{c}\left(\frac{b+cn}{b+c}\right)(k-z) \\
&\quad -\frac{K}{n}\left(q_G G_- - \widetilde{R} - \frac{(b+cn)^2}{bc(b+c)}(k-z) + \overline{y}n + \frac{m}{c}\right) - \frac{bc(b+c)\sigma^2}{2(b+cn)^2}.
\end{aligned}
$$

With short-term agreements,

$$
\begin{aligned}
y_i &= \overline{y}_i - \frac{nm}{b} - \frac{nc}{b}\left(\frac{b\overline{y}n - mn^2 + b\left(\widetilde{G} - R\right)}{b+cn^2}\right) = \overline{y} + \frac{b\overline{y} - mn - cn\left(\widetilde{G} - R\right)}{b+cn^2} \text{ and} \\
g_i &= \overline{y}_i - \overline{y} + \frac{b\overline{y} - mn - cn\left(\widetilde{G} - R\right)}{b+cn^2} - R_i.
\end{aligned}
$$

$$
\begin{aligned}
G &= \frac{k}{c} - \frac{m+z}{c} + \frac{b\theta}{b+cn^2} \Rightarrow \\
\overline{y} - \widetilde{y}_i &= \frac{nm}{b} + \frac{nc}{b}\left(\frac{k}{c} - \frac{m+z}{c} + \frac{b\theta}{b+cn^2}\right) = \frac{n}{b}\left(k - z + \frac{bc\theta}{b+cn^2}\right) \Rightarrow \\
u_i^{st} &= -\frac{c}{2}G^2 - \frac{b}{2}\left(\overline{y} - \widetilde{y}_i\right)^2 - Kr \\
&= -\frac{c}{2}\left(\frac{k}{c} - \frac{m+z}{c} + \frac{\theta b}{b+cn^2}\right)^2 - \frac{n^2}{2b}\left(k - z + \frac{\theta bc}{b+cn^2}\right)^2 - Kr \Rightarrow \\
\mathrm{E}u_i^{st} &= -\frac{1}{2}(k-z)^2\left(\frac{1}{c} + \frac{n^2}{b}\right) - \frac{m^2}{2c} + \frac{m(k-z)}{c} \\
&\quad -\frac{K}{n}\left(q_G G_- - q_R R_- + n\overline{y} + \frac{m}{c} - \left(\frac{b+cn^2}{b}\right)\left(\frac{k}{c} - \frac{z}{c}\right)\right) - \frac{\sigma^2 bc}{2\left(b+cn^2\right)}.
\end{aligned}
\tag{8.11}
$$

Comparing (8.7) with (8.10) and (8.5) with (8.11),

$$
\begin{aligned}
R^{no} - R^{st} &= -\frac{k\left(b+nc\right)^2}{bc\left(b+c\right)} + \frac{z\left(b+nc\right)^2}{bc\left(b+c\right)} + \left(\frac{b+cn^2}{b}\right)\left(\frac{k}{c} - \frac{z}{c}\right) \\
&= \frac{k\left(n-1\right)^2}{b+c}\left(1 - \frac{\delta q_R K}{nk}\right) > 0. \\
G^{no} - \mathrm{E}G^{st} &= \left(\frac{k}{c} - \frac{z}{c}\right)\left(\frac{b+nc}{b+c} - 1\right) = k\left(\frac{n-1}{b+c}\right)\left(1 - \frac{\delta q_R K}{nk}\right) = \frac{R^{no} - R^{st}}{n-1} > 0.
\end{aligned}
$$

$$
\begin{aligned}
\mathrm{E}u_i^{st} - \mathrm{E}u_i^{no} &= -\frac{1}{2}(k-z)^2\left(\frac{1}{c} + \frac{n^2}{b} - \left(\frac{1}{c} + \frac{1}{b}\right)\left(\frac{b+nc}{b+c}\right)^2\right) + m\frac{k-z}{c}\left(1 - \frac{b+nc}{b+c}\right) \\
&\quad -\frac{K}{n}(k-z)\left(\frac{\left(b+nc\right)^2}{bc\left(b+c\right)} - \frac{b+cn^2}{bc}\right) + \frac{\sigma^2 bc}{2}\left(\frac{b+c}{\left(b+cn\right)^2} - \frac{1}{b+cn^2}\right) \\
&= \left(\frac{\sigma^2 bc}{2\left(b+nc\right)^2\left(b+cn^2\right)} - \frac{(k-z)^2}{2bc\left(b+c\right)} + \frac{K}{nbc\left(b+c\right)}(k-z)\right) \\
&\quad \bullet \left[(b+c)\left(b+cn^2\right) - \left(b+nc\right)^2\right] - \frac{m(k-z)}{b+c}(n-1)
\end{aligned}
$$

$$
= \left(\frac{\left(bc\sigma\left[n-1\right]\right)^2}{2\left(b+nc\right)^2\left(b+cn^2\right)} - \frac{(k-z)\left[n-1\right]^2}{2\left(b+c\right)}\left[k - z - \frac{2K}{n}\right]\right) - \frac{m(k-z)(n-1)}{(b+c)}.
$$

Thus, we get $U^{st} > U^{no}$ if

$$
\begin{aligned}
&\mathrm{E}u_i^{st} - \mathrm{E}u_i^{no} + m\frac{(k-z)(n-1)}{(b+c)} - z\frac{k\left(n-1\right)^2}{b+c}\left(1 - \frac{\delta q_R K}{nk}\right) = \\
&\left(\frac{\left(bc\sigma\right)^2\left[n-1\right]^2}{2\left(b+nc\right)^2\left(b+cn^2\right)} - \frac{\left[n-1\right]^2}{2\left(b+c\right)}\left[(k-z)^2 - \frac{2K}{n}(k-z) + 2zk\left(1 - \frac{\delta q_R K}{nk}\right)\right]\right) > 0
\end{aligned}
$$

$$\Rightarrow \frac{(bc\sigma)^2 (b+c)}{(b+nc)^2 (b+cn^2)} > \left(\frac{K}{n}\right)^2 \left[\left(n\frac{1-\phi}{\mu}-1\right)^2 - (1-\delta q_R)^2\right].$$

Proof of Proposition 4.
In the last period, investments are given by:

$$\begin{aligned} k &= B'(g_{i,T}+R_{i,T}-\overline{y}_i)+z \Rightarrow \\ \widetilde{y}_i - \overline{y} &= -\frac{k-z}{b},\ R_{i,T}=\overline{y}_i - g_{i,T} - \frac{k-z}{b}, \\ r_i &= \overline{y}_i - g_{i,T} - \frac{k-z}{b} - q_R R_{i,-}. \end{aligned} \tag{8.12}$$

Anticipating the equilibrium $R_{i,T}$, i can invest q_R less units in period T for each invested unit in period $T-1$. Thus, in period $T-1$, equilibrium investments are given by:[25]

$$\begin{aligned} k &= B'(g_{i,T-1}+R_{i,T-1}-\overline{y}_i)+\delta q_R K \Rightarrow \\ R_{i,T-1} &= \overline{y}_i - g_{i,T-1} - \frac{k-\delta q_R K}{b}, \\ r_{i,T-1} &= \overline{y}_i - g_{i,T-1} - \frac{k-\delta q_R K}{b} - q_R R_{i,-}. \end{aligned}$$

The same argument applies to every period $T-t$, $t \in \{1,...T-1\}$, and the investment level is given by the analogous equation for each period but T. Proposition 4 follows since the optimal R_i and r_i, given g_i, are:

$$\begin{aligned} R_i^* &= \overline{y}_i - g_i - \frac{K(1-\delta q_R)}{b}, \\ r_i^* &= \overline{y}_i - g_i - \frac{K(1-\delta q_R)}{b} - q_R R_{i,-}. \end{aligned}$$

Proof of Proposition 5.
If the negotiations fail, the default outcome is the noncooperative outcome, giving everyone the same utility. Since the r_is follow from the g_is in (8.12), everyone understands that negotiating the g_is is equivalent to negotiating the r_is. All countries have identical preferences w.r.t. the r_is (and their default utility is the same), and symmetry requires that r_i, and thus $\varsigma_t \equiv \overline{y}_i - g_{i,t} - q_R R_{i,t-1}$, is the same for all countries, in equilibrium.

For the last period, (8.12) becomes

$$r_{i,T} = \varsigma_T - \frac{k - \delta q_R K/n}{b}.$$

Anticipating the equilibrium investments, the utility for the last period is:

$$U_i = -\frac{(k-z)^2}{2b} - \mathrm{E}C(G) - K r_{i,T} + \delta U(G,R).$$

[25] This presumes that country i's cost of future technology is K, which is correct since, in equilibrium, country i pays $K(1-\phi)$ plus the subsidy ϕK (or minus the tax $-\phi K$), even if this price is for the remaining fraction μ, after the fraction $1-\mu$ has been expropriated.

Efficiency requires (f.o.c. of U_i w.r.t. ς recognizing $g_i = \overline{y}_i - q_R R_{i,-} - \varsigma$ and $\partial r_i / \partial \varsigma = 1 \forall i$):

$$n\mathrm{E}C'(G) - K - n\delta U_G + n\delta U_R = 0 \Rightarrow \mathrm{E}C'(G) + m + z = K/n. \quad (8.13)$$

For the earlier periods, $t < T$, $r_{i,t} = r_{j,t} = r_t$ and

$$r_t = \varsigma_t - \frac{k - \delta q_R K}{b}.$$

Note that for every $t \in (1, T)$, $R_{i,t-1}$ is given by the g_i in the previous period:

$$\begin{aligned} r_t &= \left(\overline{y}_i - g_{i,t} - q_R \left(\overline{y}_i - g_{i,t-1} - \frac{k - \delta q_R K}{b}\right)\right) - \frac{k - \delta q_R K}{b} \\ &= \overline{y}_i (1 - q_R) - g_{i,t} + q_R g_{i,t-1} - (1 - q_R) \frac{k - \delta q_R K}{b}. \end{aligned} \quad (8.14)$$

All countries have the same preferences over the ς_ts. Dynamic efficiency requires that the countries are not better off after a change in the ς_ts (and thus the $g_{i,t}$s), given by $(\Delta\varsigma_t, \Delta\varsigma_{t+1})$, such that G is unchanged after two periods, i.e., $\Delta\varsigma_{t+1} = -\Delta\varsigma_t q_G$, $t \in [1, T-1]$. From (8.14), this implies

$$\begin{aligned} -n\mathrm{E}C'(G_t)\Delta\varsigma_t + \Delta g_t K + \delta(\Delta\varsigma_{t+1} - \Delta g_t q_R) K - \delta^2 \Delta g_{t+1} q_R K &\leq 0 \forall \Delta\varsigma_t \Rightarrow \\ \left(-\mathrm{E}C' n + K - \delta(q_G + q_R) K + \delta^2 q_G d_R K\right)\Delta\varsigma_t &\leq 0 \forall \Delta\varsigma_t \Rightarrow \\ (1 - \delta q_R)(1 - \delta q_G) \frac{K}{cn} &= \mathrm{E}G = \mathrm{E}G^*. \end{aligned}$$

Thus, neither G_t nor $g_{i,t}$ (and, hence, neither R) can be functions of R_-. At the start of period 1, therefore, $U_{R_-} = q_R K/n$, just as before, and U_G cannot be a function of R (since $U_{RG} = 0$). Since $\mathrm{E}G$ is a constant, we must have $\varsigma_1 = \overline{y} - (\mathrm{E}G^* - q_G G_0)/n - q_R R_0 / n$. (8.12) gives $\partial r_{i,t=1}/\partial G_- = (\partial r_i/\partial g_i)(\partial g_i/\partial\varsigma)(\partial\varsigma/\partial G_-) = q_G/n$. Hence, $U_{G_-} = -q_G K/n + \delta U_R q_G = -q_G(1 - \delta q_R) K/n$, giving a unique equilibrium. Substituted in (8.13), $\mathrm{E}G_T = \mathrm{E}G^*$, just as in the earlier periods. Thus, $g_{i,t}^{lt} = g_i^*(r_i^*)$ in all periods.

Proposition 5 follows since, from (3.1), $\partial g_i^*/\partial r_j = -b/(b + cn^2)$, so $g_{i,t}^{lt} = g_i^*(r_i^*) = g_i^*(r_{i,t}^{lt}) - (r_i^* - r_{i,t}^{lt}) b/(b + cn^2)$.

Proof of Proposition 6.
The optimal T balances the cost of underinvestment when T is short and the cost of the uncertain θ, increasing in T. In period T, countries invest suboptimally not only because of the domestic hold-up problem, but also because of the international one. When all countries invest less, u_i declines. The loss in period T, compared to the earlier periods, is:

$$\begin{aligned} H &= B(y_{i,t} - \overline{y}_i) - B(y_{i,T} - \overline{y}_i) - K(r_{i,t} - r_{i,T})(1 - \delta q_R) \\ &= -\frac{b}{2}\left(\frac{k - \delta q_R K}{b}\right)^2 + \frac{b}{2}\left(\frac{k - z}{b}\right)^2 - K\left(\frac{k - z}{b} - \frac{k - \delta q_R K}{b}\right)(1 - \delta q_R) \\ &= \delta q_R K^2 \frac{1 - 1/n}{b}\left[\frac{1 - \phi}{\mu} + 3\delta q_R \frac{1 - 1/n}{2}\right]. \end{aligned}$$

Note that H increases in n, q_R, K, but decreases in μ, ϕ, and b.

The cost of a longer-term agreement is associated with θ. Although $\mathrm{E}C'$ and thus $\mathrm{E}G_t$ is the same for all periods,

$$\begin{aligned}
\mathrm{E}\frac{c}{2}\left(G_t\right)^2 &= \mathrm{E}\frac{c}{2}\left(\mathrm{E}G_t+\sum_{t'=1}^{t}\theta_{t'}q_G^{t-t'}\right)^2=\frac{c}{2}\left(\mathrm{E}G_t\right)^2+\mathrm{E}\frac{c}{2}\left(\sum_{t'=1}^{t}\theta_{t'}q_G^{t-t'}\right)^2\\
&= \frac{c}{2}\left(\mathrm{E}G_t\right)^2+\frac{c}{2}\sigma^2\sum_{t'=1}^{t}q_G^{2(t-t')}=\frac{c}{2}\left(\mathrm{E}G_t^2\right)+\frac{c}{2}\sigma^2\left(\frac{1-q_G^{2t}}{1-q_G^2}\right).
\end{aligned}$$

For the T periods, the total present discounted value of this loss is L, given by:

$$\begin{aligned}
L(T) &= \sum_{t=1}^{T}\frac{c}{2}\sigma^2\delta^{t-1}\left(\frac{1-q_G^{2t}}{1-q_G^2}\right)=\frac{c\sigma^2}{2\left(1-q_G^2\right)}\sum_{t=1}^{T}\delta^{t-1}\left(1-q_G^{2t}\right)\\
&= \frac{c\sigma^2}{2\left(1-q_G^2\right)}\left[\frac{1-\delta^T}{1-\delta}-q_G^2\left(\frac{1-\delta^Tq_G^{2T}}{1-\delta q_G^2}\right)\right]\Rightarrow \qquad (8.15)\\
L'(T) &= \frac{c\sigma^2\left(-\delta^T\ln\delta\right)}{2\left(1-q_G^2\right)}\left[\frac{1}{1-\delta}-\frac{q_G^{2T+2}\left(1+\ln\left(q_G^2\right)/\ln\delta\right)}{1-\delta q_G^2}\right].
\end{aligned}$$

If all future agreements last $\widehat{T}$ periods, the optimal T for this agreement is given by

$$\begin{aligned}
&\min_T L(T)+\left(\delta^{T-1}H+\delta^TL\left(\widehat{T}\right)\right)\left(\sum_{\tau=0}^{\infty}\delta^{\tau\widehat{T}'}\right)\Rightarrow\\
&0=L'(T)+\delta^T\ln\delta\left(H/\delta+L\left(\widehat{T}\right)\right)=L'(T)+\delta^T\ln\delta\left(H/\delta+L\left(\widehat{T}\right)\right)\\
&= -\delta^T\ln\delta\left[\frac{c\sigma^2/2}{1-q_G^2}\left(\frac{1}{1-\delta}-\frac{q_G^{2T+2}\left(1+\ln q_G^2/\ln\delta\right)}{1-\delta q_G^2}\right)-\frac{H/\delta+L\left(\widehat{T}\right)}{1-\delta^{\widehat{T}'}}\right], \qquad (8.16)
\end{aligned}$$

assuming some T satisfies (8.16). Since $\left(-\delta^T\ln\delta\right)>0$ and the bracket-parenthesis increases in T, the loss decreases in T for small T but increases for large T, and there is a unique T minimizing the loss (even if the loss function is not necessarily globally concave). Since the history (G_- and R_-) does not enter in (8.16), T satisfying (8.16) equals $\widehat{T}$, assuming also $\widehat{T}$ is optimal. Substituting $\widehat{T}=T$ and (8.15) in (8.16) gives:

$$H/\delta=\frac{c\sigma^2q_G^2}{2\left(1-q_G^2\right)\left(1-\delta q_G^2\right)}\left(\frac{1-\delta^Tq_G^{2T}}{1-\delta^T}-q_G^{2T}\left(1+\frac{\ln\left(q_G^2\right)}{\ln\delta}\right)\right), \qquad (8.17)$$

where the r.h.s. increases in T. $T=\infty$ is optimal if the left-hand side of (8.17) is larger than the right-hand side even when $T\to\infty$:

$$\frac{c\sigma^2q_G^2}{2\left(1-q_G^2\right)\left(1-\delta q_G^2\right)}\leq H/\delta. \qquad (8.18)$$

If k/K and n are large, but b small, H is large and (8.18) is more likely to hold and if it does not, the T satisfying (8.17) is larger. If c or σ^2 are large, (8.18) is less likely to hold and if it does not, (8.17) requires T to decrease.

Proofs of Propositions 7-9.

Proposition 7 follows since $\phi > 0$ is allowed in the proofs above (k is a function of ϕ).[26] To see Proposition 8: Under short-term agreements (as well as business as usual), if interrim utility is $W\left(\widetilde{G}, R\right)$, investments are given by $\mathrm{E}W_R = k$ while they should optimally be $\mathrm{E}W_R = K/n$, requiring (5.2). For long-term agreements, investments are optimal in the last period if $k - \delta q_R K/n = K(1 - \delta q_R)$, requiring (5.3). For earlier periods, the requirement is $k = K$, giving (5.4). Proposition 9 follows from the text.

Proof of Proposition 10.

First, note that there is never any trade in permits in equilibrium. Hence, *if* country i invests as predicted in Sections 3-4, the marginal benefit of more technology is the same whether permits are tradable or not. Second, if i deviated by investing more (less), it's marginal utility of a higher technology decreases (increases) not only when permit-trade is prohibited, but also when trade is allowed since more (less) technology decreases (increases) the demand for permits and thus the equilibrium price. Hence, such a deviation is not attractive. When permits are tradable, altering their allocation is a form of side transfer, making the feasibility of explicit transfers irrelevant.

[26] Some caution is necessary, however. The proofs of Propositions 4-6 are unchanged only if the innovator receives the subsidy or pays the tariff *before* negotiating the price. With the reverse timing, ϕ would have no impact when the buyer is a government. In that case, the subsidy must be paid by foreign countries (as an international subsidy), and the proofs of Propositions 4-6 would need minor modifications, although the results would continue to hold. The proofs of Propositions 0-3 can stay unchanged in all these cases.

References

Acemoglu, Daron; Antras, Pol and Helpman, Elhanan (2007): "Contracts and Technology Adoption," *American Economic Review* 97 (3): 916-43.

Aldy, Joseph; Barrett, Scott and Stavins, Robert (2003): "Thirteen Plus One: A Comparison of Global Climate Policy Architectures," *Climate Policy* 3 (4): 373-97.

Aldy, Joseph and Stavins, Robert (Ed.) (2007): *Architectures for Agreement: Addressing Global Climate Change in the Post-Kyoto World*. Cambridge U. Press.

Barrett, Scott (2005): "The Theory of International Environmental Agreements," *Handbook of Environmental Economics 3*, edited by K.-G. Mäler and J.R. Vincent.

Başar, Tamer and Olsder, Geert Jan (1999): *Dynamic Noncooperative Game Theory.* Siam, Philadelphia.

Battaglini, Marco and Coate, Stephen (2007): "Inefficiency in Legislative Policymaking: A Dynamic Analysis," *American Economic Review* 97 (1): 118-49.

Battaglini, Marco; Nunnaro, Salvatore and Palfrey, Thomas R. (2010): "Political Institutions and the Dynamics of Public Investments," mimeo, Princeton University.

Beccherle, Julien and Tirole, Jean (2010): "Regional Initiatives and the Cost of Delaying Binding Climate Change Agreements," mimeo, Toulouse School of Economics.

Buchholz, Wolfgang and Konrad, Kai (1994): "Global Environmental Problems and the Strategic Choice of Technology," *Journal of Economics* 60 (3): 299-321.

Dockner, Engelbert J.; Jørgensen, Steffen; Van Long, Ngo and Sorger, Gerhard (2000): *Differential Games in Economics and Management Science*, Cambridge U. Press.

Doraszelski, Ulrich and Pakes, Ariel (2007): "A Framework for Applied Dynamic Analysis in IO," *Handbook of Industrial Organization* 3, North-Holland, Amsterdam: 1887-1966.

Dutta, Prajit K. and Radner, Roy (2004): "Self-enforcing climate-change treaties," *Proc. Nat. Acad. Sci. U.S.*, 101, 4746-51.

Dutta, Prajit K. and Radner, Roy (2009): "A Strategic Analysis of Global Warming: Theory and Some Numbers," *Journal of Economic Behavior & Organization* 71 (2): 187-209.

Ellman, Matthew (2006): "The Optimal Length of Contracts with Application to Outsourcing," Universitat Pompeu Fabra, DP 965.

Engwerda, Jacob C. (2005): *LQ Dynamic Optimization and Differential Games.* Wiley.

Fershtman, Chaim and Nitzan, Shmuel (1991): "Dynamic voluntary provision of public goods," *European Economic Review* 35 (5): 1057-67.

Fudenberg, Drew and Tirole, Jean (1991): *Game Theory.* MIT Press.

Gatsios, Konstantine and Karp, Larry (1992): "How Anti-Merger Laws can Reduce Investment, Help Producers, and Harm Consumers," *Journal of Industrial Economics* 40 (3): 339-48.

Golombek, Rolf and Hoel, Michael (2005): "Climate Policy under Technology Spillovers," *Environmental and Resorce Economics* 31 (2): 201-27.

Grossman, Gene M. and Helpman, Elhanan (1991): *Innovation and Growth in the Global Economy.* MIT Press.

Guriev, Sergei and Kvasov, Dmitriy (2005): "Contracting on Time," *American Economic Review* 95 (5): 1269-1385.

Harris, Milton and Holmstrom, Bengt (1987): "On The Duration of Agreements," *International Economic Review* 28 (2): 389-406.

Harstad, Bård (2010): "Incomplete Contracts in Dynamic Games," mimeo, Northwestern University.

Hart, Oliver D. and John Moore (1988). "Incomplete Contracts and Renegotiation," *Econometrica* 56: 755-85.

Hoel, Michael (1993): "Intertemporal properties of an international carbon tax," *Resource and Energy Economics* 15 (1): 51-70.

Hoel, Michael and de Zeeuw, Aart (2009): "Can a Focus on Breakthrough Technologies Improve the Performance of International Environmental Agreements?" NBER Working Paper 15043.

Houba, Harold; Sneek, Koos and Vardy, Felix (2000): "Can negotiations prevent fish wars?" *Journal of Economic Dynamics and Control* 24 (8): 1265-80.

Jaffe, Adam B.; Newell, Richard G. and Stavins, Robert N. (2003): "Technological Change and the Environment," *Handbook of Environmental Economics* 1 (edited by K.-G. Mäler and J.R. Vincent), Elsevier.

Karp, Larry S. and Zhao, Jinhua (2009): "A Proposal for the Design of the Successor to the Kyoto Protocol," in Aldy and Stavins (2009).

Kolstad, Charles D. and Toman, Michael (2005): "The Economics of Climate Policy," *Handbook of Environmental Economics* 3: 1562-93.

Levhari, David and Mirman, Leonard J. (1980): "The Great Fish War: An Example Using a Dynamic Cournot-Nash Solution," *Bell Journal of Economics* 11 (1): 322-34.

Maskin, Eric and Tirole, Jean (2001): "Markov Perfect Equilibrium: I. Observable Actions," *Journal of Economic Theory* 100 (2): 191-219.

Muuls, Mirabelle (2009): "The effect of investment on bargaining positions. Overinvestment in the case of international agreements on climate change," mimeo, Imperial College London.

Newell, Richard G.; Jaffe, Adam B. and Stavins, Robert N. (2006): "The Effects of Economic and Policy Incentives on Carbon Mitigation Technologies," *Energy Economics* 28: 563-78.

Ploeg, Frederick van der and de Zeeuw, Art (1992): "International aspects of pollution control," *Environmental and Resource Economics* 2 (2): 117-39.

Sorger, Gerhard (2006): "Recursive Nash Bargaining Over a Productive Asset," *Journal of Economic Dynamics and Control* 30 (12): 2637-59.

Yanase, Akihiko (2006): "Dynamic Voluntary Provision of Public Goods and Optimal Steady-State Subsidies," *Journal of Public Economic Theory* 8 (1): 171-9.