The Dynamics of Climate Agreements

Bård Harstad*

18 January 2010

Abstract

I develop a dynamic model of private provision of public bads allowing investments in technologies. The analysis is tractable and the MPE unique. The framework is used to derive optimal incomplete contracts in a dynamic setting. While the noncooperative equilibrium is very inefficient, short-term contracts can be worse due to hold-up problems. The optimal long-term contract is more ambitious if its length is relatively short and the technological spillover large. The optimal length increases in this externality. With renegotiation, the outcome is first best. The results have several implications for how to design a climate treaty.

Key words: Dynamic private provision of public goods, dynamic common pool problems, dynamic hold-up problems, incomplete contracts, contract-length, renegotiation design, climate change and climate agreements

JEL: D86, Q54, F55, F53, H87

* 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, Ben Jones, Larry Karp, Charles Kolstad, Matti Liski, Jiang Ning, Wolfgang Pesendorfer, and Roberto Veneziani. Judith N. Levi assisted with the editing.

1. Introduction

This paper develops a dynamic model of private provision of public goods. The agents can also invest in cost-reducing technologies, leading to $n + 1$ stocks, but the analysis is nevertheless tractable. I characterize a unique Markov perfect equilibrium (MPE), compare it to scenarios in which the agents can contract on contributions, and derive the optimal incomplete contract.

The model is general and could fit various contexts. The leading example is climate change, and the results have clear implications for how to design an efficient treaty. Consistent with the model's assumptions, environmental agreements (e.g. the Kyoto Protocol) typically specify emission levels but not investments in technology, since such investments would be hard to verify. They often have a limited time horizon and leave future commitments to be negotiated.[1] To fix ideas, I thus refer to the agents as "countries", the public bad (i.e., the negative of a public good) as "greenhouse gas" and contributions as "emissions." All countries suffer from the cumulated pollution level, but each country faces a private cost when cutting its own emission. This cost, however, can be reduced by investing in technology (such as abatement technology or renewable energy sources). There might also be technological spillovers when a country makes such investments, since other countries may be able to utilize the knowledge thereby generated.

The real investment cost function may be convex or concave (if there are increasing returns to scale). By assuming it is linear, I analytically derive a unique MPE, even though there is a large number of stocks in the model. This MPE is stationary and coincides with the unique subgame perfect equilibrium if time were finite but approached infinity. Since the MPE is unique, 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. While abstracting from domestic politics, I vary the countries' possibilities or negotiating, contracting and committing, and derive the best agreement

[1] According to the UN, "The major feature of the Kyoto Protocol is that it sets binding targets...for reducing greenhouse gas (GHG) emissions... over the five-year period 2008-2012...By the end of the first commitment period of the Kyoto Protocol in 2012, a new international framework needs to have been negotiated" (http://unfccc.int/kyoto_protocol/items/2830.php). The more recent Copenhagen Accord also requests the specification of emission levels, but not of levels of R&D.

for each situation. Since each equilibrium contract 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, in the future, induced to pollute less since the problem is then more severe. They will also invest more in technology to be able to afford the necessary cuts in emissions. On the other hand, 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 emission levels: 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. This makes everyone worse off, particularly if the length of an agreement is short and the number of countries large.

Long-term agreements may better mitigate 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 under-invest if (i) the agreement does not last forever or (ii) the technological spillover is large. To encourage more investments, the best (and equilibrium) long-term agreement is tougher and stipulates lower emissions compared to the optimum ex post, particularly if the technological spillover is large and the length of the agreement relatively short. Since investments decrease toward the end of the agreement, the agreement should become tougher over time to motivate investments. The optimal length of an agreement increases in the technological spillover, I find.

However, such long-term agreements are not renegotiation-proof. Once the invest-

[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, p. A4). Although the Kyoto Protocol aimed for uniform cuts relative to the 1990 levels, exceptions were widespread and there is currently no attempt to harmonize cuts.

ments are sunk and the state of the world realized, countries have an incentive to negotiate ex-post optimal emission levels rather than sticking to an overambitious long-term agreement. When renegotiation is possible and cannot be prevented, an investing country understands that it does not, in the end, have to comply with overambitious contracts. Nevertheless, with renegotiation, all investments and emissions are *first best*. Intuitively, emission levels are renegotiated to ex-post optimal levels. Countries with poor technology find it particularly costly to comply with an initial ambitious agreement and will be quite desperate to renegotiate it. This gives them a weak bargaining position and a bad outcome. To avoid this fate, countries invest more in technology, particularly if the initial agreement is very ambitious. Taking advantage of this effect, the agreement should be tougher if its length is short and the technological spillover large, just as in the case without renegotiation.

The next section clarifies the paper's contribution to the literature on dynamic games, incomplete contracts, and environmental agreements. The model is presented in Section 3. When solving the model in Section 4, I gradually increase the possibilities for negotiations and contracts: (i) no cooperation, (ii) short-term agreements, (iii) long-term agreements, and (iv) long-term agreements with renegotiation. I start out by assuming that investments are noncontractible, technologies nontradable, quotas nontradable, firms nonexistent, and countries homogeneous. All these assumptions are relaxed in Section 5. Section 6 concludes, while the Appendix contains all proofs.

2. Contributions to the Literature

By developing a (i) dynamic (difference) game permitting (ii) incomplete contracts interpreted as (iii) environmental agreements, the paper contributes to three strands of literature.

2.1. Dynamic games

Private provision of public goods is often studied in differential games (or a difference game, if time is discrete) where each player's action influences the future stock or state parameter.[3] Given the emphasis on stocks, the natural equilibrium concept is Markov perfect equilibrium. As in this paper, the typical conclusion is that public bads (goods) are over-provided (under-provided).[4]

Differential games are, however, often difficult to analyze. This has several implications: First, many authors restrict attention to linear-quadratic functional forms (Engwerda, 2005). Second, while some papers arbitrarily select the linear MPE (e.g., Fershtman and Nitzan, 1991), multiple equilibria typically exist (Tutsui and Mino, 1990). Consequently, many scholars, like Dutta and Radner (2009), manage to construct more efficient nonlinear MPEs. Third, few bother complicating their model further by adding investments in technologies. One exception is Dutta and Radner (2004), who do add explicit investments in technology. But since the cost of pollution (as well as the cost of R&D) is assumed to be linear, the equilibrium is "bang-bang" where countries invest either zero or maximally in the first period, and never thereafter.

The first contribution of this paper is to develop a tractable model, with a unique MPE, that can be used to analyze investments as well as emissions. This is achieved by assuming that technology has a linear cost and an additive impact. This trick simplifies the model tremendously, and it may also be used when studying unrelated topics. In the literature on industry dynamics, for example, analytical solutions are rare and numerical simulations necessary.[5]

My second contribution, made possible by the first, is to incorporate incomplete contracts in dynamic games. Few papers allow for policies or negotiation in differential games.[6] In Battaglini and Coate (2007), legislators negotiate spendings on pork and a

[3] For overviews, see Başar and Olsder (1999) or Dockner et al. (2000).

[4] This follows if private provisions are strategic substitutes (as in Fershtman and Nitzan, 1991, and Levhari and Mirman, 1980). If they were complements, e.g. due to a discrete public project, efficiency is more easily obtained (Marx and Matthews, 2000).

[5] See the survey by Doraszelski and Pakes (2007). A firm typically over-invests in capacity to get a competitive advantage. While Reynolds (1987) restricts attention to the linear MPE in a linear-quadratic model, a simple two-stage game is used by d'Aspremont and Jacquemin (1988) to discuss the benefits of cooperation and by Gatsios and Karp (1992) to show that firms may invest more if they anticipate future merger negotiations. When allowing negotiations on price (but not on investments) in a more general setting, Fershtman and Pakes (2000) use numerical analysis.

[6] 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) study one-period agreements. Although Ploeg and de Zeeuw (1992) even allow for R&D, contracts are either absent or complete in all these papers.

long-lasting public good. The equilibrium public-good level is suboptimally but strategically low to discourage future coalitions from wasting money on pork. This argument relies on majority rule, however, and the contract incompleteness is related to future policies rather than current investments.

2.2. Incomplete contracts

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).[7] I show that investments decrease toward the end of a contract, the best contract becomes tougher over time, and the optimal length increases in the spillover.[8] The result that short-term agreements can be worse than no agreement at all is certainly at odds with the classical literature that focuses on bilateral trade.

When renegotiation is possible, moral hazard problems are often expected to worsen (Fudenberg and Tirole, 1990). But Chung (1991) and Aghion et al. (1994) have shown how the initial contract can provide incentives by affecting the bargaining position associated with particular investments. While these (and related) models have only one period, Guriev and Kvasov (2005) present a dynamic moral hazard problem emphasizing the termination time. Their contract is renegotiated at every point in time, to keep the remaining time horizon constant. Contribution levels are not negotiated, but contracting on time is quite similar to contracting on quantity, as studied by Edlin and Reichelstein (1996): if the externality increases, Guriev and Kvasov find that the contract length should increase, while Edlin and Reichelstein show that the contracted quantity should increase. In this paper, agents can contract on quantity (emissions) as well as on time, allowing me to study how the two interact. I also allow an arbitrary number of agents, in contrast to the buyer-seller situations in these papers.

[7] In dynamic settings, hold-up problems may be solved if the parties can invest while negotiating and agreements can be made only once (Che and Sakovics, 2004), or if there are multiple equilibria in the continuation game (Evans, 2008). Neither assumption is satisfied in this paper, however.

[8] Very few papers study the optimal length of contracts. Harris and Holmstrom (1987) discuss length when contracts are costly to rewrite, but uncertainty about the future makes it necessary. Ellman (2006) studies the contract "length" (the probability for continuing the contract) and finds that it should last longer if specific investments are important. This is related to my result on the optimal time horizon, but Ellman has only two agents, one investment period, and uncertainty is not revealed over time.

2.3. Environmental agreements

There is a growing literature on climate policy and environmental agreements.[9] My main contribution to this literature is to add dynamics and incomplete contracts. This generates several novel results, including my finding that short-term agreements are bad while long-term agreements better mitigate hold-up problems. Karp and Zhao (2009), for example, recommend decade-long short-term agreements, partly to ensure flexibility. The present paper demonstrates, however, that flexibility is better assured by long-term agreements with renegotiation.

Assuming nonverifiable R&D is quite standard.[10] Thus, the result that agreements should be ambitious in order to induce R&D has been observed also in two-stage games (Golombek and Hoel, 2005). But my result that (short-term) agreements can *reduce* R&D is at odds with most of the literature, which instead emphasizes the positive impacts of regulation on technological change.[11] That R&D might decrease prior to negotiations was first noted by Buchholtz and Konrad (1994). More recently, Beccherle and Tirole (2010) generalized my one-period model and showed that anticipating negotiations can have adverse effects also if the countries, instead of investing, sell permits on the forward market, allow banking, or set production standards. With only one period, however, these models miss dynamic effects and thus consequences for agreement design.

3. The Model

3.1. Stocks and Preferences

This section presents a game where n players over time contribute to a public good and invest in technology. The purpose of the technology is to reduce the cost of providing public goods in the future. While the model is general, let climate change fix ideas. I will thus refer to the players as countries, the public good (or its negative: the public bad) as

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

[10] If trying to contract on R&D, Golombek and Hoel (2005, p. 202) note that "it will be relatively easy for the country to have less R&D than required by the agreement, but to report other expenditures as R&D activities."

[11] See, e.g., Jaffe et al. (2003) or Newell et al. (2006).

the stock of greenhouse gases, and to contributions as emissions.[12]

The public bad is represented by the stock G of greenhouse gases in *excess* of 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 increase, nevertheless, depending on the emission level g_i from country $i \in \{1, ..., n\}$:

$$G = q_G G_- + \theta + \sum_i g_i. \tag{3.1}$$

G_- represents the stock of greenhouse gases left from the previous period; subscripts for periods are thus skipped. The shock θ, arbitrarily distributed with mean 0 and variance σ^2, captures Nature's stochastic impact on G. I abstract from country-specific uncertainty.

The other type of stock is technology. The technology stock in country i is measured by R_i, and it depreciates over time at the rate $1 - q_R \in [0, 1]$. If country i invests r_i units in technology, R_i increases directly by dr_i units and, allowing for technological spillovers, R_j may increase by er_i, $\forall j \neq i$.[13]

$$R_i = q_R R_{i,-} + dr_i + \sum_{j \neq i} er_j. \tag{3.2}$$

Assuming the spillover is imperfect, $d > e$. The total effect on $R \equiv \sum_i R_i$ is defined by the primitive constant $D \equiv d + e(n-1)$.

With only one type of technology, I cannot distinguish among innovation, development, diffusion, and learning by doing. Thus, several 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:

$$g_i = y_i - R_i. \tag{3.3}$$

Alternatively, R_i may measure the capacity of country i's renewable energy sources (or "windmills"). If the windmills can generate R_i units of energy, and the alternative is to

[12]Nevertheless, I abstract from heterogeneities within countries and oil exhaustability. The strategic effects studied below would survive if these complications were added to the model.

[13]Such spillovers are empirically important (Coe and Helpman, 1995).

use polluting fossil fuels, the total amount of energy produced is given by $y_i = g_i + R_i \Rightarrow$ (3.3).

Relying on (3.3), rather than focusing on technologies that reduce the emission *content* of *each* produced unit (e.g., $g_i = y_i/R_i$), simplifies the analysis tremendously. An equally helpful assumption is to let the investment cost be linear and equal to Kr_i. In reality, the cost of investing in technology can be a convex or a concave function (if there are increasing returns to scale). Imposing linearity is thus a strong assumption, but it permits a tractable model despite the $n+1$ stocks.

Let the benefit of consuming (and producing) energy be given by the increasing and concave function $B(y_i)$. If $C(G)$ is an increasing convex function representing each country's cost of the public bad, i's utility in a period is:

$$u_i = B(y_i) - C(G) - Kr_i.$$

Remarks on θ: The stochastic shock θ has a minor role in the model and most of the results hold without it (i.e., if $\sigma = 0$). But the future marginal cost of emissions is in reality uncertain, and this can be captured by θ. In fact, the model would be identical if the level of greenhouse gases were $\widehat{G} \equiv q_G G_- + \sum_i g_i$ and the uncertainty were in the associated cost-function, affecting C but not $\widehat{G}$:

$$u_i = B(y_i) - C\left(\widehat{G} + \Theta\right) - Kr_i, \text{ where } \Theta = q_G \Theta_- + \theta.$$

Most results would continue to hold even if the effects of $\widehat{G}$ and Θ were not additive.[14] Note that, although θ is i.i.d. across periods, it has a long-lasting impact through its effect on G.

Alternative interpretations: Instead of interpreting y_i as "energy," we could substitute (3.3) in $B(.)$ and let $B(g_i + R_i)$ measure i's direct benefit of adding to the public bad (e.g., due to saved abatement costs). Better technology is then a perfect substitute for producing the public bad. The public bad does not, of course, have to be greenhouse gases. Moreover, by defining a public *good* as $-G$ and contributions as $-g_i$, i's marginal cost of providing the public good is $B'(R_i - (-g_i))$, increasing in i's contribution but decreasing

[14] The exceptions are Propositions 3 and 6 where I rely on quadratic utility functions.

in i's technology. Hence, the model fits many situations (with private provision of public goods or bads) beyond climate change.

3.2. The Timing

The investment stages and the pollution stages alternate over time. Somewhat arbitrarily, I define "a period" to be such that the countries first (simultaneously) invest in technology, after which they (simultaneously) decide how much to pollute. In between, θ is realized. Information is symmetric at all stages.

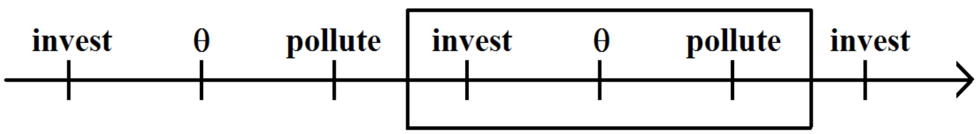

Figure 1: The timing and definition of a period

Country i's objective is to maximize the present-discounted value of its future utilities,

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

where δ is the common discount factor and U_i is i's continuation value as measured at the start of period t. As mentioned, subscripts denoting period t are often skipped.

3.3. The Equilibrium Concept

As in most dynamic games with stocks, attention is restricted to Markov perfect equilibria (MPE) as defined by Maskin and Tirole (2001). The MPE turns out to be unique and coinciding with the unique subgame-perfect equilibrium if time were finite and approaching infinity.[15] Maskin and Tirole (2001, pp. 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

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

capturing that "bygones are bygones more completely than does the concept of subgame-perfect equilibrium." While this rules out trigger strategies and related punishments, I will nevertheless consider the possibility that countries can negotiate future emission levels. Although I do not explain *why* countries comply with such promises, one explanation is that the treaty must be ratified domestically and that certain stakeholders have incentives to sue the government unless it complies. By increasing the possibilities for negotiating and contracting, I derive results for each situation, making a comparison feasible.

If the countries are negotiating, I assume the outcome is efficient and symmetric *if* the payoff-relevant variables are symmetric across countries. These assumptions are weak and satisfied whether we rely on (i) the Nash Bargaining Solution (with or without side transfers) or (ii) take-it-or-leave-it offers (with side transfers) if each country has the same chance of being recognized as the proposer. All countries participate in equilibrium, since there is no stage at which they can commit to not negotiate with the others.

4. Analysis

This section solves the game above, gradually increasing the possibilities for negotiating and contracting. As a reference for the future, the first-best emission level g_i^* ex post (taking R, G_-, and θ as given) satisfies

$$B' - nC' + n\delta U_G = 0, \text{ where} \tag{4.1}$$
$$B' \equiv \partial B\left(g_i^* + R_i\right)/\partial g_i,\ C' \equiv \partial C\left(G\right)/\partial G,\ U_G = -q_G\left(1 - \delta q_R\right)K/Dn.$$

The first-best investment level is given by

$$\mathrm{E}B'\left(g_i + R_i^*\right) = \frac{K\left(1 - \delta q_R\right)}{D} \Rightarrow \tag{4.2}$$

$$\mathrm{E}C'\left(G\right) = \frac{\left(1 - \delta q_G\right)\left(1 - \delta q_R\right)K}{Dn}. \tag{4.3}$$

Expectations are w.r.t. the unknown realization of θ. Combined with (3.1), (4.3) pins down the $\sum_i g_i$s. Given the g_is, (4.2) determines the first-best R_i^*s which, with (3.2), determine the first-best r_is. Throughout the analysis, I assume $g_i \geq 0$ and $r_i \geq 0$ never bind.[16]

[16] This is satisfied if $g_i < 0$ and $r_i < 0$ are allowed, or if q_G and q_R are sufficiently small.

4.1. No Agreement

First, assume that the countries act noncooperatively at every stage. This may be reasonable if the countries cannot commit to future policies because effective sanctions are lacking.

Note that choosing g_i is equivalent to choosing y_i, once the R_is are sunk. Substituting (3.3) into (3.1), we get:

$$G = q_G G_- + \theta + \sum_i y_i - R, \tag{4.4}$$

$$R \equiv \sum_i R_i = q_R R_- + \sum_i r_i D. \tag{4.5}$$

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

At the emission stage, when the technologies are sunk, i solves

$$\max_{y_i} B(y_i) - C(G) + \delta U(G, R) \text{ s.t. } (4.4) \Rightarrow$$

$$B' - C' + \delta U_G = 0. \tag{4.6}$$

First, note that, since (4.6) decreases in g_i, each country pollutes too much compared to the first best (4.1) when $U_G < 0$: a country is not internalizing the cost for the others.

Second, (4.6) confirms that each i chooses the same y_i, no matter the R_is. While perhaps surprising at first, the intuition is straightforward. Every country has the same preference (and marginal utility) w.r.t. y_i, and the *marginal* impact on G is also the same for every country: one *more* energy unit generates one unit of emissions.[18]

Substituting (4.4) in (4.6) implies that a larger R 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. Furthermore, substituting (3.3) in (4.6) implies that if R_i increases, g_i must decrease. In sum: if country i has better technology, i pollutes less but (because of this) other countries pollute

[17] This follows from the definition by Maskin and Tirole (2001, p. 202), where Markov strategies are measurable with respect to the coarsest partition of histories consistent with rationality.

[18] This follows from (3.3) and would be false if technology affected the marginal emission content of energy production.

more. Clearly, this effect discourages countries from investing. In addition, the Appendix shows that, in equilibrium, r_i increases in G_- but decreases in R_-. Thus, a country may want to pollute a lot and invest little today in order to induce the other countries to invest more tomorrow. With these dynamic effects, this common pool problem is more severe than its static counterpart (or the open-loop equilibrium).[19]

Proposition 1: *There is a unique symmetric MPE. Countries pollute too much and invest too little. Furthermore:*

$$\begin{aligned} y_i^{no} &= y_j^{no} \; \forall i,j \in \{1,...n\} \forall R_i, R_j, \\ \partial g_i^{no}/\partial R_i &< 0 < \partial g_i^{no}/\partial R_j \; \forall j \neq i, \end{aligned}$$

$$\partial r_i^{no}/\partial G_- = q_G/Dn, \qquad \partial r_i^{no}/\partial R_- = q_R/Dn,$$

$$U_R^{no} = q_R K/Dn, \qquad U_G^{no} = -q_G\left(1-\delta q_R\right)K/Dn.$$

Conveniently, the continuation value U is linear in G_- and R_-. Thus, the $n+1$ stocks collapse to one, making the analysis tractable. This is thanks to (3.3) and the linear investment cost, which also ensures that the equilibrium is unique.[20] Note that the equilibrium is also stationary.

The dynamic paths are simple. Following a large θ, every country pollutes less and, in the next period, investments increase such that G_+ returns to the original level. The steady state is thus reached in one period.

The size of the externality e has no effect, given D. Since only R matters for utilities and strategies, R_i becomes a pure public good.[21]

For a given R, countries pollute more if q_R is large and K/D small. Intuitively, if the technology is efficient, cheap, and long-lasting, one can pollute more today and let this motivate investments in technology tomorrow. This, and other comparative statics can

[19] This is also the case in Ploeg and de Zeeuw (1991), for example.

[20] 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. Multiple equilibria never arise under long-term agreements.

[21] Thus, adding to the public good $-G$ (by reducing g_i) or to R (by increasing r_i) has somewhat similar effects. However, they are not equivalent since a larger r_i reduces emissions in *every* future period. Increasing r_i thus has longer-lasting impact than reducing g_i, which is why r_i is referred to as an investment. Moreover, the remainder of this section lets g_i be contractible while r_i is not.

be observed in the Appendix. The working paper version derives explicit formulae for the case where $B(.)$ and $C(.)$ are quadratic functions.

4.2. 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 for the countries to develop new technology during the time-span of the agreement and the relevant technology is given by historic investments. This interpretation of short-term agreements can be captured by the timing of Figure 2.

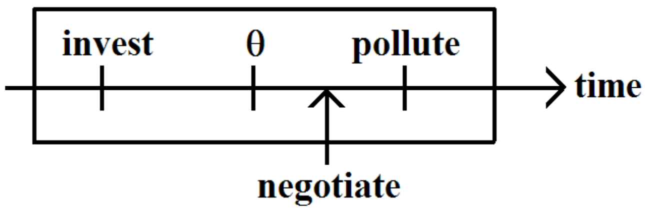

Figure 2: The timing for short-term agreements

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, (4.4)-(4.5) 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 y_i whether negotiations succeed or not. Symmetry thus implies that y_i is the same for every country in the bargaining outcome. Efficiency implies that the y_is are optimal (all countries agree on this):

$$\begin{aligned} B' - nC' + n\delta U_G &= 0 \Rightarrow \\ g_i^{st} &= g_i^*, \end{aligned} \tag{4.7}$$

where both g_i^* and g_i^{st} are functions of existing technology and pollution levels.

Substituting (4.4) in (4.7) and (3.3) in (4.7) implies that if R_i increases, g_i must decrease - but g_j increases, $\forall j \neq i$. Intuitively, if i has better technology, i's marginal benefit from polluting is less (and i is also polluting less in equilibrium). 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. The Appendix shows i invests until

$$\mathrm{E}B'\left(g_i + R_i^{st}\right) = \frac{K\left(n - \delta q_R\right)}{D}, \tag{4.8}$$

so R_i^{st} is smaller than the optimal one, given by (4.2). The equilibrium pollution level is

$$\mathrm{E}C'\left(G\right) = \frac{\left(1 - \delta q_G\right)\left(1 - \delta q_R\right)K}{Dn} + \frac{K\left(1 - 1/n\right)}{D}.$$

Thus, although emission levels are ex post optimal (4.7), once the investments are sunk, G is larger compared to its first-best level (4.3) since the hold-up problem discourages investments and makes it ex post optimal to pollute more.

PROPOSITION 2: *Proposition 1 continues to hold: There is a unique symmetric MPE. Countries pollute too much and invest too little. Furthermore:*

$$\begin{aligned}
y_i^{st} &= y_j^{st}\ \forall i, j \in \{1, ...n\} \forall R_i, R_j, \\
\partial g_i^{st}/\partial R_i &< 0 < \partial g_i^{st}/\partial R_j\ \forall j \neq i, \\
\partial r_i^{st}/\partial G_- = q_G/Dn, &\qquad \partial r_i^{st}/\partial R_- = q_R/Dn, \\
U_R^{st} = q_R K/Dn, &\qquad U_G^{st} = -q_G\left(1 - \delta q_R\right)K/Dn.
\end{aligned}$$

While its intuition is quite different, Proposition 2 is identical to Proposition 1. 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: its level can be different. But this does imply that in 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 no agreement.

4.2.1. 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 equilibrium investments and utilities?

Unfortunately, a general comparison is not feasible. But some insight can be generated by assuming $B''(.)$ and $C''(.)$ are constants:

$$B(y_i) = -\frac{b}{2}(\overline{y} - y_i)^2 \text{ and } C(G) = \frac{c}{2}G^2. \tag{Q}$$

Parameters $b > 0$ and $c > 0$ measure the importance of energy and climate change. The bliss point $\overline{y}$ represents the ideal energy level if there were no concern for pollution: no country would produce more than $\overline{y}$ due to the implicit costs of generating, transporting, and consuming energy.

Proposition 3: *Under (Q), short-term agreements reduce (i) pollution, (ii) investments, and (iii) utilities if n is large and each period short (i.e., if (4.9) holds):*

$$\begin{aligned}
\mathrm{E}G^{st} &= \mathrm{E}G^{no} - \frac{K}{D}\left(\frac{n-1}{b+c}\right)\left(1 - \frac{\delta q_R}{n}\right), \\
r_i^{st} &= r_i^{no} - \frac{K(n-1)^2}{nD^2(b+c)}\left(1 - \frac{\delta q_R}{n}\right), \\
(n-1)^2 - (1-\delta q_R)^2 &> \sigma^2\left[\frac{(b+c)(bcnD/K)^2}{(b+cn^2)(b+cn)^2}\right].
\end{aligned} \tag{4.9}$$

Rather than being encouraging, short-term agreements impair the motivation to invest. The reason is the following. Anticipating negotiations, the hold-up problem is exactly as strong as the crowding-out problem in the noncooperative equilibrium: in either case, each country only enjoys $1/n$ of the total benefits generated by its investments (no matter e). 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 invests less.[22]

Since investments decrease under short-term agreements, utilities can decrease as well. This is the case, in particular, if the period for which the agreement lasts is truly short. If so, δ and q_R are large, while there is not much uncertainty from one period to the next. All changes make (4.9) reasonable, and it always holds when the agreement is very short ($\sigma \to 0$). Moreover, (4.9) is more likely to hold if n is large (it always holds if $n \to \infty$):

[22] A counter-argument is that, if an agreement is expected, it becomes more important to invest to ensure a decent energy consumption level. While this force is smaller under (Q), it could dominate for other functional forms.

the under-investment problem is then large, it becomes important to increase investments, and this is achieved by having no agreement.

At the emission stage, however, 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 (4.9) holds, the countries would have been better off if they could commit to not negotiate short-term agreements. In particular, it may be better to commit to emission levels before the investments occur.

4.3. Long-term Agreements

The model can (and will) be used to analyze agreements of any length. If the countries can negotiate and commit to future emission levels, it will be possible to develop technologies within the time-frame of an agreement. The other countries are then unable to hold up the investing country, since the quotas have already been negotiated, at least for the near future.

4.3.1. One-period Agreements

This interpretation of "long-term agreements" can be captured simply by letting the countries negotiate the g_is in the beginning of each period, before the investments are made. While these agreements last only one period, they are indeed "longer" than the short-term agreements studied in Section 4.2. Moreover, each period can be arbitrarily long in the model, since I have not specified whether the discount factor, for example, is large or small.

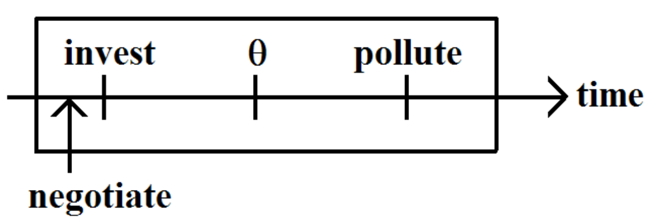

Figure 3: The timing for long-term agreements

For each period, the timing is now reversed. When investing, a country prefers a larger stock of technology if its quota, g_i^{lt}, is small, since otherwise it is going to be very costly to comply. Consequently, r_i decreases in g_i^{lt}. The Appendix shows that r_i increases until

$$B'\left(g_i^{lt}+R_i^{lt}\right)=\frac{K\left(1-\delta q_R/n\right)}{D-e\left(n-1\right)}. \tag{4.10}$$

Compared to (4.8), (4.10) suggests that countries invest more under long-term than under short-term agreements (at least for the same g_i). But compared to the first best (4.2), countries still under-invest if $e>0$ or $\delta q_R>0$. First, a country does not internalize the spillover e on the other countries. Second, if the agreement does not last forever ($\delta>0$), a country anticipates that good technology worsen its bargaining position in the future, once a new agreement is to be negotiated. At that stage, good technology leads to a lower $g_{i,+}^{lt}$ since the other countries can hold up i when it is cheap for i to reduce its emissions.[23] This discourages i from investing now, particularly if the current agreement is relatively short (δ large) and the technology likely to survive (q_R large). In sum, if e, δ, and q_R are large, it is important to encourage more investments. This can be achieved by a small g_i^{lt}.

The Appendix shows that the equilibrium and optimal g_i^{lt}s must satisfy (4.3): the equilibrium pollution level is similar to the first best! But since (4.10) implies that the equilibrium R_i^{lt}s are less than optimal, the g_i^{lt}s are suboptimally low ex post. Combining (4.3) and (4.10):

$$B'-\mathrm{E}nC'-n\delta U_G=\frac{K}{D}\left(\frac{e(1-\delta q_R)(n-1)+\delta q_R\left(1-1/n\right)}{D-e\left(n-1\right)}\right)\Rightarrow \tag{4.11}$$

$$g_i^{lt}=\mathrm{E}g_i^*-\frac{K}{D\left(b+cn^2\right)}\left(\frac{e(1-\delta q_R)(n-1)+\delta q_R\left(1-1/n\right)}{D-e\left(n-1\right)}\right)\text{ if (Q).}$$

Taking the investments as given, optimally the g_i^{lt} should have satisfied $B'-\mathrm{E}nC'-n\delta U_G=0$ rather than (4.11). Only that would equalize marginal costs and benefits of abatement. Relative to this ex post optimal level, the g_i^{lt} satisfying (4.11) must be *lower*. If e and δ are large, the right-hand side of (4.11) is large, and g_i must decline. This makes the long-term agreement more demanding or *tougher* to satisfy at the emission stage. The purpose of such an overambitious agreement is to encourage investments, since these are suboptimally low when e and δ are large.

PROPOSITION 4: *(i) There is a unique MPE. (ii) Each country invests more if the agreement is tough (4.10). Therefore, (iii) the optimal agreement (4.11) is tougher if the*

[23] Or, if no agreement is expected in the future, a large $R_{i,+}$ reduces $g_{i,+}$ and increases $g_{j,+}$, as proven in Section 4.1.

externality e is large and the time horizon short (δ large).

On the other hand, if $e = \delta q_R = 0$, the right-hand side of (4.11) is zero, meaning that the commitments under the best long-term agreement also maximize the expected utility ex post. In this case, there are no externalities, and the countries are not concerned with how current technologies affect future bargaining power. Thus, investments are first best and there is no need to distort the g_i^{lt}s downwards.

The continuation value U is linear in the stocks, making the analysis tractable. Moreover, $U_R^{lt} = q_R K/Dn$ and $U_G^{lt} = -q_G (1 - \delta q_R) K/Dn$, just as in the previous cases. 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.

4.3.2. Multiperiod Agreements

Assume now that at the beginning of period 1, countries negotiate the $g_{i,t}$s for every period $t \in \{1, 2, ..., T\}$. When investing in period $t \in \{1, 2, ..., T\}$, countries take the $g_{i,t}$s as given, and the continuation value in period $T+1$ is $U(G_T, R_T)$. At the last investment stage, i's problem is the same as before and i invests until (4.10) holds. Anticipating this, i can invest less in period T by investing more in period $T-1$. The net investment cost is thus $K(1 - \delta q_R)$. The same logic applies to every previous period and, in equilibrium,

$$\mathrm{E}B'(g_{i,t} + R_{i,t}) = \frac{K(1 - \delta q_R)}{d} = \frac{K(1 - \delta q_R)}{D - e(n-1)} \text{ for } t < T. \tag{4.12}$$

Thus, the incentives to invest are larger earlier than in the last period (4.10). In fact, if $e = 0$, investments are first best for every $t < T$. In the last period, however, countries invest less, anticipating the hold-up problem in period $T + 1$.[24]

All this is anticipated when the countries negotiate the $g_{i,t}$s. As shown in the Appendix, the optimal $g_{i,t}$s must satisfy (4.3) for every $t \le T$: the pollution level is similar to the first best! The $g_{i,t}$s are thus lower than what is optimal ex post when $e > 0$ and countries

[24] Or, if no agreement is expected in period $T+1$, i anticipates $\partial g_{j,T+1}/\partial R_i > 0$, $j \neq i$.

under-invest. Combining (4.3) and (4.12) for $t < T$,

$$B' - \mathrm{E}nC' - n\delta U_G = \frac{K}{D}\left(\frac{e(n-1)(1-\delta q_R)}{D - e(n-1)}\right) \Rightarrow \tag{4.13}$$
$$g_{i,t} = \mathrm{E}g_i^* - \frac{K}{D(b + cn^2)}\left(\frac{e(n-1)(1-\delta q_R)}{D - e(n-1)}\right) \text{ if (Q).}$$

Ex post, $B' - nC' - n\delta U_G = 0$ is optimal. Compared to this, $g_{i,t}$ satisfying (4.13) should be smaller if e is positive and large. For $t = T$, however, (4.11) continues to hold and since its right-hand side is less than that of (4.13), $y_{i,T} < y_{i,t}$ for $t < T$. In words: in order to encourage investments, the agreement should be tougher to satisfy toward the end.

Proposition 5: *Suppose countries negotiate emission levels for T periods. (i) There is a unique MPE. (ii) Investments decrease toward the end and, to encourage more investments, (iii) the equilibrium agreement becomes tougher over time compared to the ex post optimum (4.11)-(4.13).*

4.3.3. The Optimal Length of an Agreement

The optimal T trades off two concerns. On the one hand, investments are particularly low just before a new agreement is to be negotiated. This hold-up problem arises less frequently if T is large. On the other hand, the stochastic θ makes it hard to predict the optimal $g_{i,t}$s for the future, particularly when T is large. If θ were known or contractible, the agreement should last forever. Otherwise, one can show that the optimal T declines as $e \downarrow 0$. The Appendix derives a large number of comparative statics for the case where B and C are quadratic (Q):

Proposition 6: *Under (Q), the agreement's optimal length T increases in the externality e and the number of countries n but decreases in b, c, and σ.*

Intuitively, the under-investment problem is particularly severe if e and n are large. Reinforcing this problem by a small T is then especially harmful, and the optimal T is larger. Naturally, T should be smaller if future optimal emissions are uncertain (σ large) and important (c large).[25]

[25] If b is large, consuming energy is much more important than the concern for future bargaining power, the hold-up problem vanishes, and the optimal T is smaller.

4.4. Long-term Agreements with Renegotiation

The long-term agreements above are not renegotiation-proof. Not only are the commitments made before the severity of the problem (determined by θ) is known, but they also specify emission levels that are less than what is *expected* to be optimal ex post. The countries may thus be tempted to renegotiate the treaty, after θ and the investments are realized. This section derives equilibria when renegotiation is costless.

4.4.1. One-period Agreements and Renegotiation

The timing in each period is now the following. First, the countries negotiate the initial commitments, the g_i^{de}s, referred to as "the default." If these negotiations fail, it is natural to assume that the threat point is no agreement.[26] Thereafter, the countries invest and θ is realized. Before carrying out their commitments, the countries get together and renegotiate the g_i^{de}s. Relative to the threat point g_i^{de}, the bargaining surplus is split equally.[27]

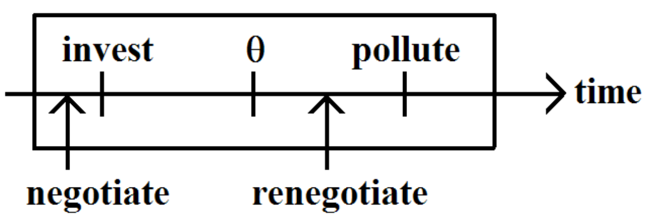

Figure 4: The timing when renegotiation is possible

Renegotiation ensures that emission levels are ex post optimal, in contrast to the long-term agreements in Section 4.3. When investing, a country anticipates that it will not, in the end, have to comply with an overambitious long-term agreement. Will this jeopardize the incentives to invest?

Proposition 7: *(i) There is a unique MPE. (ii) The initial agreement satisfies (4.14): the initial quota g_i^{de} is thus smaller if the spillover e is large and the time horizon short*

[26] If the threat point were short-term agreements, negotiated after the investment stage, the outcome would be identical.

[27] If instead the threat point at the renegotiation stage were that the countries would get upset and revert to no cooperation, the renegotiation game would be identical to negotiations under short-term agreements, and the incentives to invest would be as discussed in Section 4.2.

(δ large). (iii) All investments and emissions are first best.

$$B'\left(g_i^{de}+R_i^*\right)=\frac{K}{D-en}\Rightarrow \tag{4.14}$$
$$g_i^{de} = \mathrm{E}g_i^*-\frac{K}{bD}\left(\frac{en}{D-en}+\delta q_R\right) \text{ under (Q).}$$

When investing, the countries do anticipate that, after renegotiation, emissions will be ex post optimal, just as they were under a short-term agreement. But for the short-term agreement, countries with the poorest technology got the better deal, since these countries were quite satisfied with the (noncooperative) default outcome in which they could pollute more. This made the "technology-losers" reluctant to negotiate, giving them a better bargaining position. However, things are quite different when *re*negotiating an *ambitious* agreement. Then, the technology-losers are desperate to reach a new agreement that would replace the initial commitments. Such countries now have a poor bargaining position, and they are, in equilibrium, going to get quite a bad deal (where they must pay or accept a small g_i). Fearing this, the countries are induced to invest more, particularly if the default emission levels are small.

All this will be taken into account when negotiating the initial agreement, the g_i^{de}s. The more ambitious this agreement is, the more the countries invest. This is desirable if the countries are otherwise tempted to under-invest. Thus, the agreement should be more ambitious if e and δq_R are large. Formally, (4.14) implies that g_i^{de} decreases in e and δq_R since R_i^* is increasing in δq_R but independent of e. Intuitively, if the length of the agreement is short, countries fear that more technology today will hurt their bargaining position in the near future. They thus invest less than what is optimal, unless the agreement is more ambitious.[28]

Compared to (4.11), the initial agreement should be tougher than the optimal long-term agreement ($g_i^{de}<g_i^{lt}$). Intuitively, the long-term agreement (without renegotiation) balances the concern for investments (by reducing g_i^{lt}) and for ex post efficiency (in which g_i should be larger). The latter concern is irrelevant when renegotiation ensures ex post optimality, so the initial contract can be tougher - indeed so tough that investments are first best.

[28] No contract can help if the externality dominates the direct effect ($e \geq d$), as first pointed out by Che and Hausch (1999) and later generalized by Segal and Whinston (2002).

4.4.2. Multiple Periods and Renegotiation

The first best is implemented by any long-term agreement lasting $T \geq 1$ periods if renegotiation is possible. Suppose an agreement specifies $g_{i,t}^{de}$, $t \in \{1, ..., T\}$. Investments at $t < T$ are first best if:

$$B'\left(g_{i,t}^{de} + R_i^*\right) = \frac{K(1 - \delta q_R)}{D - en} \Rightarrow g_{i,t}^{de} = \mathrm{E}g_{i,t}^* - \frac{K}{bD}\left(\frac{1 - \delta q_R}{D/en - 1}\right) \text{ if (Q).} \tag{4.15}$$

Compared to (4.14), $g_{i,t}^{de}$ is larger when $T > 1$ than when $T = 1$ (R_i^* is independent of T). Thus, agreements lasting one period should be more ambitious than if $T > 1$, confirming the earlier finding that an agreement should be more ambitious if its length is short.

Proposition 8: *Suppose countries negotiate emission levels for $T > 1$ periods and renegotiation is possible. At $t < T$, all investment and emission levels are first best if $g_{i,t}^{de}$ is given by (4.15). T and $g_{i,t'}^{de}$, $t' > t$, are irrelevant.*[29]

4.4.3. Implementation

The optimal $g_{i,t+1}^{de}$s depend on θ_t and they must be (re)negotiated after θ_t is realized. But instead of being negotiated at the start of period $t + 1$, the $g_{i,t+1}^{de}$s may equally well be negotiated just before the emission stage in period t, since no information or individual decisions are made in between. At this time, therefore, the countries may negotiate every $g_{i,t+1}^{de} < \mathrm{E}g_{i,t+1}^*$ while simultaneously renegotiating the $g_{i,t}^{de}$s and replacing them by the optimal $g_{i,t+1}^*$s, which are expected to be larger than the $g_{i,t}^{de}$s negotiated in advance. This might be observationally equivalent to a time-inconsistent policy where the countries make ambitious plans for the future, while repeatedly backing down from promises made in the past. But rather than verifying a time-inconsistency problem, this leads to the first best.

Corollary 1: *In equilibrium, the countries repeatedly promise to pollute little in the future but when the future arrives, they relax these promises. This procedure implements the first best.*

[29] Since T and $g_{i,t'}^{de}$, $t' > t$, are irrelevant, the predictions are not sharp when renegotiation is feasible. With a small fixed cost of negotiating each $g_{i,t'}^{de}$, however, the unique optimal contract would be described by Proposition 7.

5. Generalizations and Extensions

5.1. Contractible Investments

If the countries negotiated R&D but not emission levels, deriving the best incomplete contract would require an analysis somewhat similar to that in Section 4. If both R&D and emission levels were contractible, the first best would trivially be implemented, even without renegotiation. Suppose optimal investments are implemented by international subsidies. Let s be a subsidy to i for each invested unit, and assume every $j \neq i$ pays $sr_i/(n-1)$ to i. Utilities become:

$$\begin{aligned} u_i &= B(y_i) - C(G) - kr_i - \sum_{j \neq i} sr_j/(n-1), \text{ where} \\ k &\equiv K - s. \end{aligned} \tag{5.1}$$

Proposition 9: *The investment subsidy should be larger if the agreement's length is short: the optimal s is given by (5.2) for short-term agreements, (5.3) for long-term agreements (and the last period of multiperiod agreements), and by (5.4) for multiperiod agreements (except for the last period).*

$$s^{st} = \frac{K(n-1)}{n} > \tag{5.2}$$

$$s^{lt} = \frac{K(n-1)}{n}\left(\delta q_R + \frac{en}{D}(1-\delta q_R)\right) > \tag{5.3}$$

$$s^{t} = \frac{Ke(n-1)}{D}. \tag{5.4}$$

5.2. Tradable Technologies

The model is tractable and can be extended in several directions. As an alternative or additional externality, let j benefit *directly* by ϵr_i when $i \neq j$ invests.[30] The model is unchanged if just s in (5.1) is replaced by $-\epsilon(n-1)$ and K continues to measure the social net cost of investing while k is the private cost. One can interpret ϵ as a *general* technological spillover (affecting u_i and not only i's environmental technology) or as a

[30] There are several ways in which spillovers could be formalized. For example, the spillover could be related to R_i rather than r_i (as in Coe and Helpman, 1995). The results would be similar, but i must then consider the impact of r_i on R_j not only for the present, but for all future periods.

spillover that reduces i's *cost* of making a particular investment r_i (such that the cost is $kr_i - \sum_{j \neq i} \epsilon r_j$).[31]

In reality, ϵ may reflect international law: Suppose r_i has the *potential* to reduce j's cost (or to increase u_j) by $\bar{\epsilon} r_i$ units. Of this, j can copy a fraction $\gamma \in [0,1]$ for free. The remaining fraction, $1-\gamma$, is available if j pays i for transferring (or licensing) its technology. Investor i may be the government or a firm acting on its behalf (outsourcing). If i sets the price and the foreign individuals or firms purchasing the benefit $(1-\gamma)\bar{\epsilon}$ face an *ad valorem* tariff λ, i charges their willingness to pay, $(1-\lambda)(1-\gamma)\bar{\epsilon}$, for each invested unit. The net externality for country j when i invests becomes

$$\epsilon = \gamma \bar{\epsilon} + \lambda (1-\gamma) \bar{\epsilon}.$$

The Appendix allows for both e and ϵ and finds them to play similar roles.

PROPOSITION 10: *With high tariffs (λ large) and weak intellectual property rights (γ large), the net externality (ϵ) is large and the optimal agreement is tougher and more long-lasting.*

Intuitively, with high tariffs and weak property-right protection, investors do not capture the large benefit experienced by foreigners and they invest less. To encourage more investments, the equilibrium agreement is tougher and more long-lasting. Since s and $-\epsilon(n-1)$ play similar roles, Proposition 9 implies that, optimally, property rights should be stronger and tariffs lower if the agreement is short-term, while Proposition 10 implies that if the subsidy s is given and small, the agreement should be tougher and longer-lasting.[32]

5.3. Tradable Permits

To simplify intuition and the reasoning, it has been assumed that quotas cannot be traded. This is not a critical assumption, however.

[31] In fact, the cost would take exactly this form if countries simultaneously choose their *targets* for the R_is and let the *expenditures* (the r_is) follow residually from (3.2) rather than vice versa.

[32] The working paper version elaborates on these extensions.

Proposition 11: *(i) All results survive with tradable permits, no matter whether side payments are available. (ii) The equilibrium and optimal permit price is B', thus increasing in e and larger if $T = 1$ than if $T > 1$.*

$B'(g_i + R_i)$ is the value of being allowed to pollute one more unit, keeping G and R constant. Proposition 11 follows by noting that, first, there is never any trade in permits in equilibrium. Hence, if i invests as above, the marginal benefit of more technology is the same. 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.[33,34]

5.4. Heterogeneity

So far, countries have been completely symmetric and there has been no heterogeneity. It did turn out, however, that for a given R_-, differences in $R_{i,-}$ (such as $R_{i,-} - R_{j,-}$) were payoff-irrelevant. It is therefore not necessary to assume that all countries start out with the same technology.

Moreover, since the continuation values are linear in R, countries are risk-neutral in that it would not matter if q_R were random, as long as the *expected* depreciation rate is $1 - q_R$. The realized depreciation can also be different for every country, as long as the expected depreciation rate is $1 - q_R$ for everyone.

A strong assumption has been that all countries had identical preferences. With quadratic utility functions, for example, it is reasonable to assume that countries have different bliss points ($\overline{y}_i$) for energy consumption. Generalizing the quadratic specification, we may write the benefit function as $B(y_i - \overline{y}_i)$, where $\overline{y}_i$ is a country-specific reference

[33] In a two-stage model, also Golombek and Hoel (2005) find that the permit price should be higher than "the Pigouvian" level to induce R&D when there are spillovers.

[34] An earlier version of this paper analyzed emission taxes and derived similar results: for example, the first best is feasible with renegotiation if the initial tax is higher than what is expected to be optimal ex post, particularly if the spillover is large and the agreement's length relatively short.

point (not necessarily bliss). Recognizing the importance of such heterogeneity, all proofs allow the reference point $\overline{y}_i$ to vary. While a large $\overline{y}_i$ increases the equilibrium g_i, the comparative statics are unchanged.[35]

PROPOSITION 12: *All results continue to hold with heterogenous bliss or reference points in* $B(y_i - \overline{y}_i)$.

6. Conclusions

This paper presents a novel dynamic game where n players contribute to a public bad while also investing in cost-reducing technologies. By assuming linear investment costs, the Markov perfect equilibrium (MPE) is unique, the continuation value linear, and the analysis tractable, despite the $1+n$ stocks. While the unique equilibrium rules out self-enforcing agreements, the framework can be employed to analyze incomplete contracts in a dynamic setting. I derive the optimal contract as a function of its length, and the optimal length is discussed. When renegotiation is possible, I characterize contracts implementing the first best.

While the method and the model are general, the assumptions are motivated by climate change and the implications for a treaty are important. First, agreements are not necessarily good. In particular, one should be careful when recommending short-term agreements since investments may fall. Second, an agreement should be more ambitious if its length is short. Although perhaps counterintuitive at first, this is required to motivate investments in new technology. Third, the agreement should be tougher and longer-term if the technological spillover is large. Since spillovers are particularly large if

[35] 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} or, if renegotiation is possible, g_i^{de}. Such small g_is would not be necessary (or optimal) for noninvesting countries. Naturally, the investing countries would require some compensation to accept the optimal emission targets. At the same time, a small g_i would not motivate i to invest if i were allowed to purchase permits from noninvesting countries with higher g_js. Thus, with heterogeneity in investment costs, it matters a great deal whether side transfers are possible and permits tradable: Proposition 11 would be false if such heterogeneity were introduced. Evaluating political instruments under heterogeneity is thus an important task for future research.

intellectual property rights are weak and tariffs large, treaties for climate and trade might be strategic substitutes. Fourth, if R&D can be subsidized internationally, the optimal subsidy is larger if the agreement is short-term. Finally, flexibility regarding future emission levels are best ensured by renegotiating long-term agreements rather than by letting them expire quickly. Since the commitments under the Kyoto Protocol expire in 2012, the current default is no agreement at all. On the other hand, when the Doha-round trade negotiations broke down, countries did not revert to the noncooperative equilibrium but to the existing set of trade agreements. The procedure used for trade is more efficient than the one currently used for climate, according to the above analysis.

The results hold whether side transfers are available, permits tradable, firms included, and whether the technology can be traded, taxed, or subsidized. Nevertheless, this paper is only one step toward a better understanding of good environmental agreements. I have not distinguished between technological innovation and diffusion, and I have abstracted from domestic politics, heterogeneity, private information, monitoring, compliance, coalition formation and the possibility of opting out of the agreement. Relaxing these assumptions are the natural next steps.

7. Appendix

All propositions are here proven with the generalizations discussed in Section 5: the value of y_i is given by the increasing and concave function $\beta(y_i - \overline{y}_i)$, countries can have different reference points $\overline{y}_i$, and r_i generates a direct externality $x = \epsilon - s/(n-1)$ on $j \neq i$ in addition to the technological spillover e:

$$u_i = \beta(y_i - \overline{y}_i) - C(G) - kr_i + x \sum_{j \neq i} r_j.$$

In Sections 3 and 4, $B(y_i) \equiv \beta(y_i - \overline{y})$ since $\overline{y}_i = \overline{y}$, and $x = 0 \Rightarrow k = K$.

While U_i is the continuation value just before 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 $\overline{y}$ is the average $\overline{y}_i$. Note that by substitution,

$$\begin{aligned} G &= \widetilde{G} + \sum_i y_i - \sum_i R_i = q_g G_- + \sum_i \widetilde{y}_i - R, \text{ and} \\ u_i &= B(y_i - \overline{y}_i) - C(G) - kr_i + x\sum r_j = B(\widetilde{y}_i - \overline{y}) - C(G) - kr_i + x\sum r_j. \end{aligned}$$

All i's are identical w.r.t. $\widetilde{y}_i$. The game is thus symmetric, no matter differences in R_i or $\overline{y}_i$, and the payoff relevant states are G and R. Analyzing the symmetric equilibrium (where symmetric countries invest identical amounts), I drop the subscript for i on U and W. The proof for the first best (4.1)-(4.3) is omitted since it would follow the same lines as the following proof.

7.1. Proof of Proposition 1

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

$$\begin{aligned} 0 &= \beta'(y_i - \overline{y}_i) - C'(G) + \delta U_G(G, R) \\ &= \beta'(\widetilde{y}_i - \overline{y}) - C'\left(\widetilde{G} - R + \sum \widetilde{y}_i\right) + \delta U_G(\widetilde{G} - R + \sum \widetilde{y}_i, R), \end{aligned} \tag{7.1}$$

implying that all $\widetilde{y}_i$s are identical and implicit functions of $\widetilde{G}$ and R only. At the investment stage, i maximizes:

$$\mathrm{E}W(\widetilde{G}, R) - kr_i = \mathrm{E}W\left(q_G G_- + \theta, \widetilde{R} + \sum_i Dr_i\right) - kr_i, \tag{7.2}$$

implying that R is going to be a function of G_-, given implicitly by $\mathrm{E}\partial W(q_G G_- + \theta, R)/\partial R = k/D$ and explicitly by, say, $R(G_-)$. In the symmetric equilibrium, each country invests $(R(G_-) - q_R R_-)/Dn$. Thus:

$$\begin{aligned} U(G_-, R_-) &= \mathrm{E}W(q_G G_- + \theta, R(G_-)) - (k - (n-1)x)\left(\frac{R(G_-) - q_R R_-}{Dn}\right) \Rightarrow \\ z/\delta &\equiv \frac{\partial U}{\partial R_-} = \frac{q_R K}{Dn} \end{aligned} \tag{7.3}$$

in every period. Hence, $U_{RG} = U_{GR} = 0$, m and U_G cannot be functions of R and (7.1) implies that $\widetilde{y}_i$, G and thus $\beta(\widetilde{y}_i - \overline{y}) - C(G) \equiv \gamma(.)$ are functions of $\widetilde{G} - R$ only. Hence, write $G\left(\widetilde{G} - R\right)$. (7.2) rewritten:

$$\mathrm{E}\gamma(q_G G_- + \theta - R) - kr_i + \delta U(G(q_G G_- + \theta - R), R)$$

and because U_R is a constant, maximizing this w.r.t. r_i makes $q_G G_- - R$ a constant, say ξ. This gives $\partial r_i/\partial G_- = q_G/Dn$ and U becomes:

$$\begin{aligned} U(G_-, R_-) &= \mathrm{E}\gamma(\xi + \theta) - Kr + \delta U(G(\xi + \theta), R) \\ &= \mathrm{E}\gamma(\xi + \theta) - K\left(\frac{q_G G_- - \xi - q_R R_-}{Dn}\right) + \delta U(G(\xi + \theta), q_G G_- - \xi) \Rightarrow \\ m/\delta &= \partial U/\partial G_- = -K\left(\frac{q_G}{Dn}\right) + \delta U_R q_G = -\frac{K q_G}{Dn}(1 - \delta q_R), \end{aligned} \tag{7.4}$$

since $G(\xi + \theta)$ and $\gamma(.)$ are *not* functions of G_- when $q_G G_- - R = \xi$. Since U_G is a constant, (7.1) implies that if R increases, $\widetilde{y}_i$ increases but G must decrease. Thus, $\partial\widetilde{y}_i/\partial R \in (0,1)$, so $\partial g_i/\partial R_j = \partial(\widetilde{y}_i - \overline{y}_i + \overline{y} - R_i)/\partial R_j > 0$ if $i \neq j$ and < 0 if $i = j$.

7.2. 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 (7.1) requires:

$$0 = \beta'(\widetilde{y}_i - \overline{y})/n - C'\left(\widetilde{G} - R + \sum \widetilde{y}_i\right) + \delta U_G(\widetilde{G} - R + \sum \widetilde{y}_i, R). \tag{7.5}$$

The rest of the previous proof continues to hold: R will be a function of G_- only, so $U_{R_-} = q_R K/Dn$. This makes $\mathrm{E}\widetilde{G} - R$ a constant and $U_{G_-} = -q_G(1 - \delta q_R)K/Dn$, 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 envelope theorem can be used to calculate equilibrium investments:

$$\begin{aligned}
\max_{r_i} \mathrm{E}\frac{1}{n}\left[\max_{\{\widetilde{y}_j\}} \sum_j \beta\left(\widetilde{y}_j - \overline{y}\right) - C\left(G\right) + \delta U\left(G,R\right)\right] - kr_i &\Rightarrow \\
\mathrm{E}C'\left(G\right)D - \mathrm{E}\delta U_G D + \mathrm{E}\delta U_R D - k &= \\
\mathrm{E}C'\left(G\right)D - \left(1-\delta q_G\right)\left(1-\delta q_R\right)K/n - \left(K+xn\right)\left(1-1/n\right) &= 0.
\end{aligned}$$

Combined with (7.5),

$$\begin{aligned}
\frac{\left(1-\delta q_G\right)\left(1-\delta q_R\right)K}{Dn} + \frac{\left(K+xn\right)\left(1-1/n\right)}{D} &= \frac{\mathrm{E}\beta'\left(\widetilde{y}_i - \overline{y}\right)}{n} + \delta U_G \Rightarrow \\
\frac{\left(n-\delta q_R\right)K}{D} + \frac{n\left(n-1\right)x}{D} &= \frac{\mathrm{E}\beta'\left(\widetilde{y}_i - \overline{y}\right)}{n}.
\end{aligned}$$

7.3. Proof of Proposition 3

The proof is omitted since it is not instructive and since it is available in the working paper version (and thus on request).

7.4. Proof of Proposition 4

When g_i is already negotiated, i invests until

$$\begin{aligned}
k &= \beta'\left(g_i + R_i - \overline{y}_i\right)d + zD \Rightarrow && (7.6)\\
\widetilde{y}_i - \overline{y} &= \beta'^{-1}\left(\frac{k-zD}{d}\right),\ R_i = \beta'^{-1}\left(\frac{k-zD}{d}\right) + \overline{y}_i - g_i, \\
dr_i &= \beta'^{-1}\left(\frac{k-zD}{d}\right) + \overline{y}_i - g_i - q_R R_{i,-} - \sum_{j\neq i} er_j. && (7.7)
\end{aligned}$$

Anticipating this, the utility before investing is:

$$U_i = \beta\left(\beta'^{-1}\left(\frac{k-zD}{d}\right)\right) - \mathrm{E}C\left(G\right) - kr_i + \sum_{j\neq i} xr_j + \delta U\left(G,R\right).$$

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 (7.7), everyone understands that negotiating the g_is is equivalent to negotiating the r_is. Since all countries have identical preferences w.r.t. the r_is (and their default utility is the same) the r_is are going to be equal for every i. Symmetry requires that r_i, and thus $\zeta \equiv [g_i + q_R R_{i,-} - \overline{y}_i]$, is the same for all countries. (7.7) becomes

$$Dr_i = \beta'^{-1}\left(\frac{k - zD}{d}\right) - \zeta.$$

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

$$\begin{aligned} -n\mathrm{E}C'(G) + K/D + n\delta U_G - nD\delta U_R(1/D) &= 0 \Rightarrow \\ \mathrm{E}C'(G) + m + z &= K/Dn. \end{aligned} \tag{7.8}$$

Combined with (7.7), neither G nor R can be functions of R_- (R_i in (7.7) and (7.8) are not functions of R_-). Thus, $U_{R_-} = q_R K/Dn$, just as before, and U_G cannot be a function of R (since $U_{RG} = 0$). (7.8) then implies that EG is a constant and, since we must have $\zeta = (\mathrm{E}G - q_G G_-)/n + q_R R_-/n - \overline{y}$, (7.7) gives $\partial r_i/\partial G_- = (\partial r_i/\partial g_i)(\partial g_i/\partial\zeta)(\partial\zeta/\partial G_-) = q_G/Dn$. Hence, $U_{G_-} = -q_G K/Dn + \delta U_R q_G = -q_G(1-\delta q_R)K/Dn$, giving a unique equilibrium, (7.3) and (7.4), just as before. Substituted in (7.8):

$$\mathrm{E}C'(G) = (1-\delta q_G)(1-\delta q_R)K/Dn. \tag{7.9}$$

This is the same pollution level as in the first best (4.3). But investments might be suboptimally low. Combining (7.9) with (7.6),

$$\beta'(g_i + R_i - \overline{y}_i)/n - \mathrm{E}C'(G) - m = \frac{1}{n}\left(\frac{k}{d} - \frac{K}{D}\right) + \frac{\delta q_R K}{Dn}\left(1 - \frac{D}{dn}\right) =$$

$$\frac{K}{Dn}\left(\frac{1+(n-1)x/K}{1-(n-1)e/D} - 1 + \delta q_R\left(\frac{(D-en)(n-1)}{Dn-en(n-1)}\right)\right) = \frac{K}{Dn}\left(\frac{xD/K + e + \delta q_R(D/n - e)}{D/(n-1) - e}\right).$$

Under (Q), $\beta' = b(g_i + R_i - \overline{y}_i)$ and $\mathrm{E}C' = c(q_G G_- + \sum_i g_i)$, so

$$\begin{aligned} (\beta'/n - \mathrm{E}C')^{lt} - (\beta'/n - \mathrm{E}C')^* &= b\left(-g_i^{lt} + g_i^*\right)/n - cn\left(g_i^{lt} - g_i^*\right) \Rightarrow \\ \mathrm{E}g_i^* - g_i^{lt} &= \frac{K/D}{b + cn^2}\left(\frac{x/K + e/D + \delta q_R(1/n - e/D)}{1/(n-1) + e/D}\right). \end{aligned}$$

7.5. Proof of Proposition 5

At the start of $t = 1$, countries negotiate emission levels for every period $t \in \{1, ..., T\}$. The investment level in period T is (7.7) for the same reasons as given above.

Anticipating the equilibrium $R_{i,T}$ (and $R_{j,T}$) i can invest q_R less units in period T for each invested unit in period $T-1$. Thus, in period $T-1$, i invests until:

$$\begin{aligned} k &= d\beta' \left(g_{i,T-1} + R_{i,T-1} - \overline{y}_i\right) + \delta q_R k \Rightarrow \qquad (7.10) \\ R_{i,T-1} &= q_R R_{i,T-1} + dr_{i,T-1} + \sum_{j \neq i} er_{j,T-1} = \overline{y}_i - g_{i,T-1} - \beta'^{-1} \left(k \left(1 - \delta q_R\right)/d\right). \end{aligned}$$

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.

In equilibrium, all countries enjoy the same $y_i - \overline{y}_i$ and default utilities. Thus, just as before, they will negotiate the g_is such that that they will all face the same cost of investment in equilibrium. Thus, $r_i = r_j = r$ and

$$Dr = \left(\overline{y}_i - g_i - q_R R_{i,t-1}\right) - \beta'^{-1} \left(k \left(1 - \delta q_R\right)/d\right).$$

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

$$\begin{aligned} Dr &= \left(\overline{y}_i - g_i - q_R \left(\overline{y}_i - g_i^{t-1} - \beta'^{-1} \left(k \left(1 - \delta q_R\right)/d\right)\right)\right) - \beta'^{-1} \left(k \left(1 - \delta q_R\right)/d\right) \\ &= \overline{y}_i \left(1 - q_R\right) - g_i + q_R g_i^{t-1} - \left(1 - q_R\right) \beta'^{-1} \left(k \left(1 - \delta q_R\right)/d\right). \qquad (7.11) \end{aligned}$$

Since $r_i = r_j$, (7.10) implies that the equilibrium $g_{i,t} + q_R R_{i,t-1} - \overline{y}_{i,t}$ is the same (say ς_t) for all is:

$$g_{i,t} + q_R^{t-\tau} R_{i,\tau-1} - \overline{y}_i = \varsigma_t, \; t \in [1, T].$$

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., $\Delta\varsigma_t q_G = -\Delta\varsigma_{t+1}$, $t \in [1, T-1]$. From (7.11), this implies

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

Using (7.10),

$$\beta' - \mathrm{E}C'(G)\,n - nm = \frac{k\,(1-\delta q_R)}{d} - (1-\delta q_R)(1-\delta q_G)\,K/D - \frac{\delta q_G\,(1-\delta q_R)\,K}{D}$$

$$= \frac{k\,(1-\delta q_R)}{d} - (1-\delta q_R)\,K/D = \left(\frac{k}{d}-\frac{K}{D}\right)(1-\delta q_R) = \frac{K}{D}\left(\frac{x/K+e/D}{1/(n-1)-e/D}\right)(1-\delta q_R)\,.$$

The $g_{i,T}$ satisfies (7.9) for the same reasons as in the previous proof (and since they do not influence any $R_{i,t}$, $t<T$). It is easy to check that U_R and U_G are the same as before.

Under (Q), $\beta' = b\,(\overline{y}_i - g_i - R_i)$, $\mathrm{E}C' = c\left(\mathrm{E}\widetilde{G}+\sum g_j\right)$, and since $\beta' - \mathrm{E}cGn - n\delta U_G = 0$ for $g_i^*(R)$, we have

$$\begin{aligned}
& b\,(\overline{y}_i - g_i - R_i) - nc\left(\widetilde{G}+\sum g_j\right) - n\delta U_G \\
-&\left[b\,(\overline{y}_i - g_i^* - R_i) - nc\left(\widetilde{G}+\sum g_j^*\right) - n\delta U_G\right] \quad = \quad \left(\frac{k}{d}-\frac{K}{D}\right)(1-\delta q_R) \Rightarrow \\
&\left(g_i^* - g_i^{lt}\right)\left(b+cn^2\right) = \left(\frac{k}{d}-\frac{K}{D}\right)(1-\delta q_R) \quad = \quad \frac{K}{D}\left(\frac{x/K+e/D}{1/(n-1)-e/D}\right)(1-\delta q_R)\,.
\end{aligned}$$

7.6. Proof of Proposition 6

The optimal T balances the cost of under-investment and the cost of not knowing future θs. In period T, countries invest suboptimally not only because of e and x, but because of the hold-up problem: one more unit of R_i in period $T+1$ is not worth much to i, since the other countries will take advantage of it and pollute more. When all countries invest less, u_i declines. The loss in period T, compared to the earlier periods, is under (Q):

$$\begin{aligned}
H &= \beta\,(y_{i,t}-\overline{y}_i) - \beta\,(y_{i,T}-\overline{y}_i) - K\,(r_{i,t}-r_{i,T})\,(1-\delta q_R) \\
&= -\frac{b}{2}\left(\frac{k\,(1-\delta q_R)}{bd}\right)^2 + \frac{b}{2}\left(\frac{k-zD}{bd}\right)^2 - \frac{K}{D}\left(-\frac{k\,(1-\delta q_R)}{bd}+\frac{k-zD}{bd}\right)(1-\delta q_R) \\
&= \frac{\delta q_R K^2}{bD^2}\left(\frac{k}{K}-\frac{1}{n}\right)\left[\left(\frac{k/K-d/D}{d^2/D^2}\right)\left(1-\frac{\delta q_R}{2}\right)+\frac{\delta q_R\,(d/D-1/n)}{2d^2/D^2}\right].
\end{aligned}$$

Note that H increases in e (for given D), x (for given K) and n but decreases in b.

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

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

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

$$
\begin{aligned}
L(T) &= \sum_{t=1}^{T} \frac{c}{2}\sigma^2 \delta^{t-1}\left(\frac{1-q_G^{2t}}{1-q_G^2}\right) = \frac{c\sigma^2}{2\left(1-q_G^2\right)}\sum_{t=1}^{T}\delta^{t-1}\left(1-q_G^{2t}\right) \\
&= \frac{c\sigma^2}{2\left(1-q_G^2\right)}\left[\frac{1-\delta^T}{1-\delta} - q_G^2\left(\frac{1-\delta^T q_G^{2T}}{1-\delta q_G^2}\right)\right] \Rightarrow \qquad (7.12) \\
L'(T) &= \frac{c\sigma^2\left(-\delta^T \ln\delta\right)}{2\left(1-q_G^2\right)}\left[\frac{1}{1-\delta} - \frac{q_G^{2T+2}\left(1+\ln\left(q_G^2\right)/\ln\delta\right)}{1-\delta q_G^2}\right].
\end{aligned}
$$

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

$$
\begin{aligned}
&\min_T L(T) + \left(\delta^{T-1}H + \delta^T L\left(\widehat{T}\right)\right)\left(\sum_{\tau=0}^{\infty}\delta^{\tau\widehat{T}'}\right) \Rightarrow \\
&0 = L'(T) + \delta^T \ln\delta\left(H/\delta + L\left(\widehat{T}\right)\right) = L'(T) + \delta^T\ln\delta\left(H/\delta + L\left(\widehat{T}\right)\right) \\
&= -\delta^T\ln\delta\left[\frac{c\sigma^2}{2\left(1-q_G^2\right)}\left(\frac{1}{1-\delta} - \frac{q_G^{2T+2}\left(1+\ln\left(q_G^2\right)/\ln\delta\right)}{1-\delta q_G^2}\right) - \frac{H/\delta + L\left(\widehat{T}\right)}{1-\delta^{\widehat{T}'}}\right] \qquad (7.13)
\end{aligned}
$$

assuming some T satisfies (7.13). 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 (7.13), T satisfying (7.13) equals $\widehat{T}$, assuming also $\widehat{T}$ is optimal. Substituting $\widehat{T} = T$ and (7.12) in (7.13) gives:

$$
H/\delta = \frac{c\sigma^2 q_G^2}{2\left(1-q_G^2\right)\left(1-\delta q_G^2\right)}\left(\frac{1-\delta^T q_G^{2T}}{1-\delta^T} - q_G^{2T}\left(1+\frac{\ln\left(q_G^2\right)}{\ln\delta}\right)\right), \qquad (7.14)
$$

increasing in T. $T = \infty$ is optimal if the left-hand side of (7.14) is larger than the right-hand side even when $T \to \infty$:

$$
\frac{c\sigma^2 q_G^2}{2\left(1-q_G^2\right)\left(1-\delta q_G^2\right)} < H/\delta. \qquad (7.15)
$$

If e (for given D), x (for given K) and n are large, but b small, H is large and (7.15) is more likely to hold and if it does not, the T satisfying (7.14) is larger. If c or σ^2 are large, (7.15) is less likely to hold and if it does not, (7.14) requires T to decrease.

7.7. Proof of Proposition 7

In the default outcome, a country's (interrim) utility is:

$$
W_i^{de} = \beta\left(g_i^{de} + R_i - \overline{y}_i\right) - C\left(\widetilde{G} + \sum g_j^{de}\right) + \delta U.
$$

Since i gets $1/n$ of the renegotiation-surplus, in addition, i's utility is:

$$W_i^{de} + \frac{1}{n}\sum_j \left(W_j^{re} - W_j^{de}\right) - kr_i + x\sum_{j\neq i} r_j, \tag{7.16}$$

where W_i^{re} is the utilities after renegotiation. Maximizing the expectation of this expression w.r.t. r_i gives the f.o.c.

$$\begin{aligned} k &= d\beta'\left(g_i^{de} + R_i - \overline{y}_i\right)(1 - 1/n) + Dz(1 - 1/n) \\ &\quad + \mathrm{E}\frac{D}{n}\partial\left(\sum W_i^{re}\right)/\partial R - \sum_{j\neq i}\frac{1}{n}\left(e\beta'\left(g_j^{de} + R_j - \overline{y}_j\right) + Dz\right). \end{aligned} \tag{7.17}$$

Requiring first-best investments, $\mathrm{E}D\left(\partial\left(\sum W_i^{re}\right)/\partial R\right) = K$, and since $\beta'\left(g_i^{de} + R_i - \overline{y}_i\right)$ must be the same for all is,

$$k = \beta'\left(g_i^{de} + R_i^* - \overline{y}_i\right)(d - D/n) + K/n \Rightarrow \beta'\left(g_i^{de} + R_i^* - \overline{y}_i\right) = \frac{kn - K}{dn - D}. \tag{7.18}$$

Combined with the optimum, (4.2),

$$\begin{aligned} \beta'\left(g_i^{de} + R_i^* - \overline{y}_i\right) - \mathrm{E}\beta'\left(g_i^* + R_i^* - \overline{y}_i\right) &= \frac{kn - K}{dn - D} - \frac{K}{D}(1 - \delta q_R) \\ &= \frac{K}{D}\left(\frac{x/K + e/D}{1/n - e/D} + \delta q_R\right). \end{aligned} \tag{7.19}$$

Since $y_i^{de} - \overline{y}_i$ is the same for every i in equilibrium, the bargaining game (when renegotiating the g_i^{de}s) is symmetric and the renegotiated g_i^{re}s become efficient (just as under short-term agreements). Since the first best is implemented, U_R and U_G are as before. Under (Q), $\beta'\left(g_i^{de} + R_i - \overline{y}_i\right) - \mathrm{E}\beta'\left(g_i^* + R_i - \overline{y}_i\right) = b\left(\mathrm{E}g_i^* - g_i^{lt}\right)$, so

$$\mathrm{E}g_i^* - g_i^{de} = \frac{K}{bD}\left(\frac{x/K + e/D}{1/n - e/D} + \delta q_R\right).$$

7.8. Proof of Proposition 8

Take period 1, and assume the countries renegotiate the $g_{i,1}$s only (a similar logic holds if they simultaneously renegotiate future emission levels). If $T > 1$, $R_{i,t}$ for $t = 2$ is given by $g_{i,2}$, nothwitstanding R_i^1 and whether the renegotiation over $g_{i,1}$ fails. Thus, the equilibrium first-order condition w.r.t. r_i^1 is:

$$\begin{aligned} k &= \left[d\beta'\left(g_{i,1}^{de} + R_{i,1} - \overline{y}_i\right) + \delta q_R k\right](1 - 1/n) \\ &\quad + \mathrm{E}\frac{D}{n}\partial\left(\sum W_i^{re}\right)/\partial R - \frac{1}{n}\sum_{j\neq i} e\beta'\left(g_{j,1}^{de} + R_{j,1} - \overline{y}_j\right) \end{aligned}$$

Requiring first-best investments, $\mathrm{E}D\left(\partial\left(\sum W_i^{re}\right)/\partial R\right)=K$, and since $\beta'\left(g_{i,1}^{de}+R_{i,1}-\overline{y}_i\right)$ must be the same for all is,

$$\begin{aligned} k\left[1-\delta q_R\left(1-1/n\right)\right] &= (d-e)\left[\beta'\left(g_{i,1}^{de}+R_i^*-\overline{y}_i\right)\right](1-1/n)+K/n\Rightarrow \\ \beta'\left(g_{i,1}^{de}+R_i^*-\overline{y}_i\right) &= \frac{k\left[n-\delta q_R\left(n-1\right)\right]-K}{dn-D}. \end{aligned} \tag{7.20}$$

Comparing (7.20) and (7.18) reveals that g_i^{de} is larger in the present case. Under (Q):

$$\mathrm{E}g_i^*-g_i^{de}=\frac{K}{Db}\left[(1-\delta q_R)\left(\frac{x/K+e/D}{1/n-e/D}\right)+\frac{\delta q_R x/Kn}{1/n-e/D}\right].$$

Similar argument holds for $t>1$, but since R^* depends on the latest realization of θ, $g_{i,t}^{de}$ must be renegotiated after θ is realized in period $t-1$.

7.9. Proofs of Proposition 9-10

Under short-term agreements (as well as under no agreement), if interrim utility is $W\left(\widetilde{G},R\right)$, investments are given by $\mathrm{E}W_R=k/D$ while they should optimally be $\mathrm{E}W_R=K/Dn$, requiring $K+x\left(n-1\right)=K/n\Rightarrow -x=K/n$. Under long-term agreements, the optimal R_i is given by $\beta'\left(g_i+R_i-\overline{y}_i\right)D+nzD=K$,which is the same as the equilibrium $\beta'd+zD=k$ if $-x=K\left(\delta q_R+en\left(1-\delta q_R\right)/D\right)/n$. For an agreement lasting $T>1$ periods, $R_{i,t}$, $t<T$, should be $D\beta'\left(g_{i,t}+R_i-\overline{y}_i\right)+\delta q_R K=K$, which is the same as the equilibrium $R_{i,t}$ if $K\left(1-\delta q_R\right)/D=k\left(1-\delta q_R\right)/d\Rightarrow -x=eK/D$. Proposition 10 follows since $x>0$ is allowed in the proofs above.

7.10. Proofs of Propositions 11-12

The proof of Proposition 11 is equivalent to the text following it, and thus omitted. Proposition 12 follows since heterogeneity is allowed in the proofs above.

References

Aghion, Philippe; Dewatripont, Mathias and Rey, Patrick (1994): "Renegotiation Design with Unverifiable Information," *Econometrica* 62: 257-82.

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.

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.

Che, Yeon-Koo and Hausch, Donald B. (1999): "Cooperative Investments and the Value of Contracting," *American Economic Review* 89 (1): 125-46.

Che, Yeon-Koo and Sakovics, Jozsef (2004): "A Dynamic Theory of Holdup," *Econometrica* 72 (4): 1063-1103.

Chung, Tai-Yeong (1991): "Incomplete Contracts, Specific Investment, and Risk Sharing," *Review of Economic Studies* 58 (5): 1031-42.

Coe, David T. and Helpman, Elhanan (1995): "International R&D spillovers," *European Economic Review* 39 (5): 859-87.

d'Aspremont, Claude and Jacquemin, Alexis (1988): "Cooperative and Noncooperative R&D in Duopoly with Spillovers," *American Economic Review* 78 (5): 1133-37.

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.

Edlin, Aaron S. and Reichelstein, Stefan (1996): "Hold-ups, Standard Breach Remedies, and Optimal Investment," *American Economic Review* 86 (3): 478-501.

Ellman, Matthew (2006): "The Optimal Length of Contracts with Application to Outsourcing," mimeo, UPF.

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

Evans, Robert (2008): "Simple Efficient Contracts in Complex Environments," *Econometrica* 76 (3): 459-91.

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

Fershtman, Chaim and Pakes, Ariel. (2000): "A Dynamic Oligopoly with Collusion and Price Wars," *RAND Journal of Economics* 31 (2): 207-36.

Fudenberg, Drew and Tirole, Jean (1990): "Moral Hazard and Renegotiation in Agency Contracts," *Econometrica* 58 (6): 1279-1319.

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 Resource Economics* 31 (2): 201-27.

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.

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.

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," mimeo, UC Berkeley.

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.

Marx, Leslie M. and Matthews, Steven A. (2000): "Dynamic Voluntary Contribution to a Public Project," *Review of Economic Studies* 67: 327-58.

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

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, Aart (1991): "A differential game of international pollution control," *Systems & Control Letters* 17 (6): 409-14.

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

Reynolds, Stanley S. (1987): "Capacity Investment, Preemption and Commitment in an Infinite Horizon Model," *International Economic Review* 28 (1): 69-88.

Segal, Ilya and Whinston, Michael D. (2002): "The Mirrless Approach to Mechanism Design With Renegotiation," *Econometrica* 70 (1): 1-45.

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

Tsutsui, Shunichi and Mino, Kazuo (1990): "Nonlinear Strategies in Dynamic Duopolistic Competition with Sticky Prices," *Journal of Economic Theory* 52: 136-61.

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