The Dynamics of Climate Agreements

Bård Harstad*
(harstad@kellogg.northwestern.edu)

14 October 2009

Abstract

I provide a novel dynamic model with private provision of public bads and investments in technologies. The analysis is tractable and the MPE unique. By adding incomplete contracts, I derive implications of and for international climate treaties. While the non-cooperative equilibrium is bad, short-term agreements are worse due to hold-up problems. A long-term agreement should be more ambitious if it is relatively short-lasting and the technological externality large. The length itself should increase in this externality. With renegotiation, the outcome is first best. The technological externalities are related to trade agreements, making them strategic substitutes to climate treaties.

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

JEL: D86, Q54, F55, F53, H87

* Thanks to the audiences at the 2009 ASSA meeting, the University of British Columbia, Caltech, the Chicago Fed, the Chicago Area Political Economy conference, the University of Copenhagen, Emory University, 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, Paris School of Economics, Simon Fraser University, Stockholm University and the World Bank. In addition or particular, I have benefitted from the comments of Philippe Aghion, Jeff Campbell, Yeon-Koo Che, Rolf Golombek, Michael Hoel, Ben Jones, Larry Karp, Charles Kolstad, Jiang Ning and Roberto Veneziani.

1. Introduction

This paper provides a novel dynamic model of private provision of public goods. The agents can also invest in cost-reducing technologies but, nevertheless, the Markov-Perfect Equilibrium (MPE) is unique and the analysis tractable. The non-cooperative outcome is compared to scenarios where the agents can contract on contributions (but not on investment), and the optimal contract is derived.

While the model fits a variety of contexts, the policy implications for climate agreements are particularly important. Environmental agreements (e.g. the Kyoto protocol) are typically specifying emissions but not investments in technology, since such efforts would be hard to verify. They often have a limited time horizon and future commitments remain 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 gases" 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 (abatement technology or renewable energy sources). There might also be technological spillovers when a country invests, since other countries may be able to utilize the generated knowledge.

The real cost of investments 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 it coincides with the unique sub-game 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 to negotiate, contract and commit, and derive the best agreement under alternative situations. Since each equilibrium contract is also the constrained optimum,

[1] According to the UN, "The major feature of the Kyoto Protocol is that it sets binding targets...for reducing greenhouse gas (GHG) emissions. These amount to an average of five per cent against 1990 levels 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 results can be interpreted normatively.

First, I assume countries act non-cooperatively 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 cut in emission. On the other hand, if a country invests a lot in abatement technology, it is 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 a similar static model. This dynamic common pool problem is thus particularly severe.

Nevertheless, "short-term" agreements on emission levels can make everyone *worse off*. A hold-up problem arises when the countries negotiate emission levels: if one country has a better technology, it can cut its emissions fairly cheaply, and the other countries will demand 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 time horizon of an agreement is short and the number of countries large.

The hold-up problem may be mitigated by "long-term" agreements. 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-lasting. Nevertheless, countries under-invest if (i) the agreement is not lasting forever or (ii) the technological spillover large. To encourage more investments, the best (and equilibrium) long-term agreement is tougher and stipulates *less* emission compared to the optimum ex post, particularly if the technological spillover is large and the time horizon of the agreement relatively short. The time horizon itself should increase in the spillover.

But such long-term agreements are not renegotiation-proof. Once the investments are sunk and the state of the world realized, countries have an incentive to negotiate ex post optimal emission levels rather than sticking to the over-ambitious long-term agreement.

[2] Hold-up problems are real. When negotiating climate policies in the European Union, "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" (*Financial Times*, October 17, 2008, p. A4).

If renegotiation is possible and cannot be prevented, an investing country realizes that it does not, in the end, have to comply to over-ambitious contracts. Nevertheless: with renegotiation, all investments and emissions are *first-best*. Intuitively, the emission levels are renegotiated to the ex post optimal levels. Countries with poor technologies find it particularly costly to comply to an initial ambitious agreement and they are going to be quite "desperate" when renegotiating it. This gives them a weak bargaining position and a bad deal. Anticipating this, countries invest more in technology, particularly if the initial agreement is very ambitious. Taking advantage of this, the agreement should be tougher if it is short-lasting and the technological spillover large, just as before. This way, the externality is endogeneized.

In reality, the externalities from investments are related to international trade and law. Poor protection of intellectual property rights allow countries to benefit without having to pay. If trade in abatement technology is possible, import tariffs may reduce the exporter's price and increase the externality for free-riders. International subsidies, either on investments or trade in abatement technologies, do the opposite. Thus, with small subsidies, high tariffs and poor protection of intellectual property rights, the externality is larger and countries under-invest. In these circumstances, the results suggest that the climate treaty should be tough and long-lasting. Vice versa, if the countries can only commit to short-lasting and weak climate treaties, investment subsidies, tariff reductions and intellectual property rights become more important.

The next section clarifies the paper's contribution to several strands of literature. The model is presented in Section 3. Section 4 solves the model under four scenarios where I gradually increase the possibilities to negotiate and contract: (i) no cooperation, (ii) short-term agreements, (iii) long-term agreements and (iv) long-term agreements with renegotiation. While Sections 3 and 4 let investments be non-contractible, Section 5 permits subsidies and relates the externality to trade policies. I start out by assuming symmetric countries, no firms, technologies cannot be traded, emission quotas cannot be traded, side transfers are feasible but investments non-contractible. All these assumptions are relaxed in Sections 5 and 6. The final section concludes, while Appendix contains all proofs.

2. Contribution to the Literature

By analyzing (i) climate agreements combining (ii) dynamic (difference) games with (iii) incomplete contracts, the paper contributes to all three strands of literature.

2.1. Environmental agreements

There is a growing literature on climate policy and environmental agreements.[3] 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 (2008), for example, propose short-term agreements (of 10-year length) to ensure flexibility. Flexibility is better provided by long-term agreements with renegotiation, according to the present paper.

Assuming non-verifiable R&D and additive spillovers is quite standard.[4] Thus, the result that agreements should be tough 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, instead emphasizing the positive impacts on technological change of regulation.[5] Also Hoel and de Zeeuw (2009) find that R&D can decrease if countries cooperate because they then abate even without new technology, altough there is no negotiation (and their analysis hinges on a "breakthrough technology" and binary abatement levels). Anticipating negotiations, Buchholtz and Konrad (1994) find, as I do, that R&D may decrease due to the hold-up problem. But all these models allow only one period, missing the full dynamic effects and thus the consequences for agreement design.[6]

[3] For excellent reviews, see Kolstad and Toman (2005) on climate policy and Barrett (2005) on environmental agreements. Aldy *et al.* (2003) and Aldy and Stavins (2007) discuss alternative architectures for climate agreements.

[4] If trying to contract on R&D, "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" (Golombek and Hoel, 2005, p. 202). The additive spillover is "used in most of the literature on climate policy in the context of interactions between countries and endogenous technology development" (Golombek and Hoel, 2004, p. 4).

[5] See e.g. Jaffe *et al.* (2003) or Newell *et al.* (2006). Even when investments are prior to negotiations, Muuls (2009) finds that they increase investments. In the two-stage model by Golombek and Hoel (2004), R&D and abatements are strategic complements when the cost of pollution is linear.

[6] A related literature in IO let firms invest in R&D before competing/colluding. Anticipating quantity competition, firms invest *too much* since that give them a competitive advantage, particularly if the

Other papers are truly dynamic.[7] Some of them study self-enforcing treaties by allowing trigger strategies (Barrett, 2005). In particular, Dutta and Radner (2009) study a differential game where countries pollute over time and the stock of greenhouse gas accumulates. Dutta and Radner (2004) add explicit investments in the 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 zero or maximally in the first period, and never thereafter. Moreover, there is no concern for bargaining power since there is no negotiation. Xepapadeas (1995) also allows R&D as well as emissions, but a special strategy set is imposed and no agreements allowed.[8]

2.2. Dynamic games

The paper is related to a general literature on dynamic private provision of public goods. Many of these models are differential games. A differential game (or a "difference game", if time is discrete) is a game where each player's action influences the future stock or state parameter.[9] Given the emphasis on these stocks, the natural equilibrium concept is Markov Perfect Equilibrium (defined by Maskin and Tirole, 2001). As in this paper, the conclusion is typically that public bads (goods) are over-provided (under-provided).[10] My contribution to this literature is to include R&D and incomplete contracts while providing a tractable model with a unique MPE.

The latter point is not trivial. While many authors arbitrarily select the linear MPE (e.g. Fehrstman and Nitzan, 1991), multiple equilibria often exist (Wirl, 1996, Tutsui and Mino, 1990). Consequently, many scholars try to construct more efficient nonlinear

spillovers small (d'Aspremont and Jacquemin, 1988; Leahy and Neary, 1997) Gatsios and Karp (1992) show that firms may invest more still if they anticipate merger *negotiations*, and the profit may be smaller (with lower prices) than if a merger is not allowed.

[7] Nordhaus and Yang (1996) presented an early regional emission game, but without R&D, negotiations or feedback (only the open-loop equilibrium were studied).

[8] Many dynamic models of climate treaties focus on the number of participants (see e.g. Barrett and Stavins, 2003, Rubio and Ulph, 2007, and the references therein. In my model, however, all countries participate in equilibrium since I do not allow them to commit to not negotiate with the others.

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

[10] The outcome is also worse than similar static models (or the open loop equilibrium; see e.g. Ploeg and de Zeeuw, 1991). These results arise when private provisions are strategic substitutes (see e.g. Fehrstman and Nitzan, 1991, or Levhari and Mirman, 1980, for a study of over-fishing). If they were complements, e.g. due to a discrete public project, efficiency is easier to obtain (Marx and Matthews, 2000).

MPEs,[11] although Rowat (2007) finds many of them unreasonable. Since multiple MPEs make predictions hard and institutional comparisons meaningless, it is important to have a unique MPE in the present analysis.

Just a few papers allow for policies or negotiation in differential games.[12] Hoel (1993) studies a differential game with an emission tax, while Yanase (2006) derives the optimal contribution-subsidy. Houba *et al.* (2000) study negotiations (over fish quotas) in a differential game where the agreement lasts forever, while Sorger (2006) lets the agreement last only one period. Investments or R&D are not allowed, so the contract is complete.[13,14]

2.3. Incomplete contracts

By permitting contracts on emissions, but not investments, the paper is in line with the literature on incomplete contracts (going back to Hart and Moore, 1988). My main contribution is to add dynamics (an infinite time horizon).[15] This allows me to study the optimal length of contracts,[16] and the length's interaction with the contracted quantity (of emissions).

Moral hazard problems are often expected to worsen when agents can renegotiate the contract (Fudenberg and Tirole, 1990). But Chung (1991) and Aghion *et al.* (1994) showed how the initial contract can provide incentives by affecting the bargaining position

[11] E.g. Dutta and Radner (2009), Dockner and Long (1993), Dockner and Sorger (1996) and Sorger (1998).

[12] Battaglini and Coate (2008) study a dynamic game where legislators negotiate pork and public debt. Equilibrium debt is suboptimally but strategically high since that discourages future coalitions from wasting money on pork (district-specific transfers). Also in the present paper, the level of public bad (pollution) is suboptimally large partly to induce the parties to behave well (and invest more) in the future.

[13] Ploeg and de Zeeuw (1992) do allow R&D in a differential pollution game, but incomplete contracts are nevertheless not considered.

[14] By emphasizing investments, the paper is related to the literature on difference games in IO (surveyed by Doraszelski and Pakes, 2007), where firms over-invest in capital to deter their competitors from entering (Spence, 1977), investing or producing (Reynolds, 1987, Maskin and Tirole, 1987).

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

[16] Very few papers study the optimal length of contracts. Harris and Holmstrom (1987) discuss the length when contracts are costly to rewrite, but uncertainty about the future makes it necessary. Ellman (2006) studies the contract "length" (actually; 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.

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. The contract is renegotiated at every point in time, to keep the remaining time horizon constant. Contribution levels (traded quantities) are not negotiated, but contracting on time is quite similar to contracting on quantity, 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.[17] In this paper, agents can contract on quantity (emissions) as well as time, allowing me to study how the two interact. For example, I find that if the length decreases, the quantity should be more cooperative. Moreover, I allow an arbitrary number of agents, in contrast to the buyer-seller situations in these papers.

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 has many applications, let climate change fix ideas. I will thus refer to the players as "countries", the public good (or its negative; the public bad) as the stock of greenhouse gases and contributions as emissions.[18]

The public bad is represented by the stock G of "greenhouse gases" (or CO_2) in *excess* of its natural level. Since the natural level is thus $G = 0$, G tends to approach 0 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 $i \in \{1, ..., n\}$:

[17] For the first best to be attainable, it is crucial that the externality is not dominating the direct effect of the investments. Otherwise, Che and Hausch (1999) find that the null-contract is optimal. These results are generalized and discussed by Segal and Whinston (2002).

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

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

G_- represents the stock of greenhouse gases left from the previous period (this way, I do not need subscripts for periods). 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 technological spillovers, R_j may increase by er_i, $\forall j \neq i$. Developing technology is a creative process and the knowledge generated may also be used in other countries, although the environment there might differ somewhat.[19] Assuming the spillover is imperfect, $d > e$. The total effect on $R \equiv \sum R_i$ is defined by the primitive constant $D \equiv d + e\,(n - 1)$.

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

With only one type of technology, I cannot distinguish between innovation, development, diffusion or 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 costlessly reduce (or clean) its potential emission. If energy production, measured by y_i, is generally polluting, the actual emission 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 use polluting fossil fuel, 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,

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

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:

$$u_i = B(y_i) - C(G) - Kr_i. \tag{3.4}$$

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 emission 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 $\widehat{G} \equiv q_G G_- + \sum_i g_i$ while 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 continue to hold if the effects of $\widehat{G}$ and Θ were not necessarily additive.[20] 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). A better technology is then a perfect substitute to 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 in i's technology. Hence, the model fits many situations (with private provision of public goods or bads) beyond climate change.

Assumptions and possible extensions: In order to get to the main results quickly, to simplify the analysis for the applied reader, and to keep the model consistent with alternative applications, I am starting out by assuming (i) no firms, (ii) no heterogeneity,

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

(iii) a particular type of technological spillover, (iv) no trade in technologies, (v) no trade in emission quotas, (vi) if negotiating, side transfers are available, but (vii) one cannot contract on, or subsidize, investments in technology. All these assumptions are relaxed in Sections 5 and 6 (and the Appendix) and the results are shown to be robust.

3.2. The Timing

The investment stages and the pollution stages alternate over time. Somewhat arbitrary, 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.

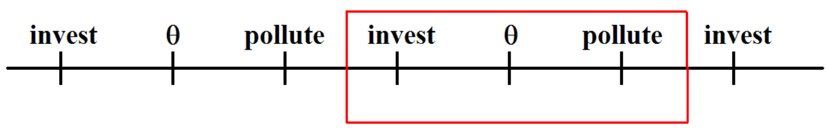

Figure 1: The timing and definition of "a period"

Since there is an infinite number of periods, country i ultimately cares about the present-discounted value of all 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 a country'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).[21] The MPE turns out to be unique

[21] For its definition, see Maskin and Tirole (2001), who also defend the MPE since it is "often quite successful in eliminating or reducing a large multiplicity of equilibria", and MPEs "prescribe the simplest form of behavior that is consistent with rationality" while capturing that "bygones are bygones more completely than does the concept of subgame-perfect equilibrium" (p. 192-3).

and coinciding with the unique subgame-perfect equilibrium if time were finite and approaching infinity.[22] While this rules out trigger strategies and related punishments, I will nevertheless consider the possibility that countries can negotiate future emission levels. I do not explain *why* countries comply to such promises, but one possibility is that the treaty must be ratified domestically and that certain stakeholders have incentives to sue the government unless it complies. By increasing the possibilities to negotiate and contract, I derive results for each situation and a comparison is feasible.

If the countries are negotiating, I assume the outcome is efficient and symmetric *if* the bargaining game is symmetric. These assumptions are quite weak, and they are 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) where each country has the same chance of being recognized as the proposer. Every country participates 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 to negotiate and contract. I first assume negotiations never take place (for example because the countries cannot commit to anything). The second subsection studies "short-term agreements" by letting countries negotiate and contract on immediate emission. Thereafter, I let countries negotiate and commit to future emission levels ("long-term agreements"), while the fourth subsection allows the countries to costlessly renegotiate such contracts. Following the incomplete contracting literature, private investments are observable but not verifiable.

As a reference for the future, the first best emission level g_i^* ex post (taking R, G_- and θ as given) is given by

$$B' - nC' + n\delta U_G = 0, \text{ where} \tag{4.1}$$

$$B' \equiv \partial B\left(g_i^* + R_i\right)/\partial g_i,\ C' \equiv \partial C\left(G\right)/\partial G,\ U_G = -q_G\left(1-\delta q_R\right)K/Dn.$$

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

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 $\sum g_i$s. Given the g_is, (4.2) determines the first best R_i^*s which, with (3.2), determines the first best r_is. Throughout the analysis, I assume $g_i \geq 0$ and $r_i \geq 0$ never binds.[23] All proofs are in the Appendix.

4.1. No Agreement

First, assume that the countries act non-cooperatively at every stage. This may be reasonable if the countries cannot commit to future policies, for example 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:

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

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

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

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

$$\max_{y_i} B\left(y_i\right)-C\left(G\right)+\delta U\left(G,R\right) \text{ s.t. } (4.4) \Rightarrow$$
$$B'-C'+\delta U_G=0. \tag{4.6}$$

[23] This is satisfied if $g_i<0$ and $r_i<0$ are allowed or q_G and q_R are sufficiently small. If the constraints exist and bind, the analysis would be rather more complicated (Wirl, 2008).

[24] 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.

First, note that 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 differences in 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. Relying on (3.3), one *more* energy unit generates one unit of emission. The abatement technology (or the windmills) is already utilized to the fullest possible extent, and producing more energy is going to pollute.

Substituting (4.4) in (4.6) implies that a larger R (by reducing G) must increase every y_i. This implies that if R_i increases but R_j, $j \neq i$, is constant, then $g_j = y_j - R_j$ must increase. Furthermore, substituting (3.3) in (4.6) implies that if R_i increases, g_i must decrease. Combined, if country i has a better technology, i pollutes less but (because of this) all other countries pollute more. Clearly, this effect discourages countries when investing, even though the total pollution declines in R_i. In addition, countries realize that if G_- is large for a given R, (3.3) and (4.6) implies that the g_is must decrease. Thus, a country may want to pollute more today to induce others to pollute less (or invest more) in the future. These dynamic considerations make this dynamic common pool problem much more severe than its static counterpart (or the open-loop equilibrium).

Proposition 1. *There is a unique symmetric MPE in which countries pollute too much, invest too little, and*

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

$$U_R^{no} = q_R K / Dn, \qquad U_G^{no} = -q_G \left(1 - \delta q_R\right) K / Dn. \tag{4.7}$$

Since U_G increases in q_R but decreases in K, (4.7) implies that countries pollute more if technology is long-lasting and cheap. For a fixed $d = D - e\,(n-1)$, a larger e increases U_G. Countries are then induced to pollute more since they benefit when the other countries, as a consequence, invest more.

Conveniently, the continuation value U is linear in G_- and R_-.[25] Thus, the $n+1$ stocks of the model basically collapses 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.[26] Note that the equilibrium is also stationary.

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-lasting, 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.

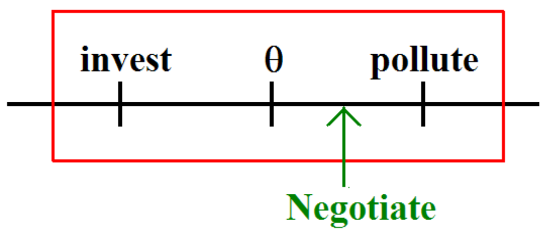

Figure 2: The timing for "short-term agreements"

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) implies 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). Since the countries are, in essence, negotiating the y_is, symmetry implies that y_i is the same for every country. Efficiency implies that the y_is are optimal (all countries agree

[25] Since the investment cost is linear, equilibrium R is a function of G_- only, not R_-. This implies that $\partial U_i^{no}/\partial R_-$ is constant (and equal to $q_R K/Dn$ in the symmetric equilibrium where all countries invest the same). For related reasons, U_i is also linear in G_-, implying that U_i is a function of the history only through its impact on $G_- - R_-\alpha$, where $\alpha \equiv q_R/q_G\,(1-\delta q_R)$ is a constant.

[26] As the proposition states, this is the unique *symmetric* MPE. There exist asymmetric MPEs, as well, in which the countries invest different amounts (since the investment cost is linear). 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 arises under long-term agreements. Section 6.2 discusses asymmetric equilibria if countries are heterogeneous.

on this):

$$
\begin{aligned}
B' - nC' + n\delta U_G &= 0 \Rightarrow \\
g_i^{st} &= g_i^*,
\end{aligned}
\tag{4.8}
$$

where g_i^* and g_i^{st} are the optimal and negotiated emission levels respectively (both functions of existing technologies and pollution).

Just as in the previous subsection, substituting (4.4) in (4.8) and thereafter (3.3) in (4.8) implies that if R_i increases, g_i must decrease but g_j increases, $\forall j \neq i$. Intuitively, if i has a good 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) reduces the marginal cost of polluting. Anticipating this hold-up problem, every country is discouraged to invest. Appendix shows that investments are such that:

$$
\mathrm{E}B'\left(g_i + R_i^{st}\right) = \frac{K\left(n - \delta q_R\right)}{D}, \tag{4.9}
$$

so the R_i^{st}s are smaller than the optimal ones given by (4.2). Combined with (4.8),

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

Thus, although emission levels are "ex post" optimal (4.8), once the investments are sunk, G is larger compared to the *first-best* (4.3) since the hold-up problem discourages investments, making it ex post optimal to pollute more.

Proposition 2. *Proposition 1 continues to hold: There is a unique symmetric MPE in which countries pollute too much, invest too little, and*

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

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 non-cooperative case. This does *not* imply that U itself is identical in the two cases: Its level can be different. But this *does* imply that when deriving actions and utilities for one period, it is irrelevant whether there will be a short-term agreement also in the next (or any future) period. This makes it convenient to compare short-term agreements to no agreement, since the comparison will be independent of whether there is an agreement in the future and whether we are referring to a one-shot agreement or a sequence of short-term agreements.

4.2.1. Are Short-Term Agreements Good?

Pollution is less under short-term agreements compared to no agreement. That should not be surprising: The entire motivation for negotiating is to reduce pollution. In fact, the emission levels are negotiated to the *first-best* emission levels given the past investments. 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, while $\overline{y}$ is the bliss point for energy production: even without any concern for pollution, a country would not produce more than $\overline{y}$ energy-units 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 (4.10) holds).*

$$\begin{aligned}
\mathrm{E}G^{st} &= \mathrm{E}G^{no} - \frac{K}{D}\left(\frac{n-1}{b+c}\right)\left(1 - \frac{\delta q_R}{n}\right) \\
r_i^{st} &= r_i^{no} - \frac{K(n-1)^2}{nD^2(b+c)}\left(1 - \frac{\delta q_R}{n}\right) \\
(n-1)^2 - (1-\delta q_R)^2 &> \sigma^2\left[\frac{(b+c)(bcnD/K)^2}{(b+cn^2)(b+cn)^2}\right] \qquad (4.10)
\end{aligned}$$

Rather than being encouraging, short-term agreements harm the motivation to invest. The reason is the following: The hold-up problem, when negotiations are anticipated, is exactly as strong as the crowding-out problem in the non-cooperative 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 country i understands that the problem will be taken care of, to some extent, since emission levels are going to be reduced. This implies that the marginal benefit of further cuts decline and it is marginally less important to invest in technologies that would further reduce future pollution. Hence, each country invests less.

Since investments decrease under short-term agreements, it may not be a surprise that utilities can decrease as well. This is the case, in particular, if the period for which the agreement lasts is quite short. If so, it is reasonable that δ is large, q_R large, while there is not much uncertainty from one period to the next. All changes make (4.10) reasonable, and it *always* holds if $\sigma = 0$. Moreover, (4.10) is more likely to hold if n is large (it *always* holds if $n \to \infty$), since then the under-investment problem is large and it is very important to increase investments. This is achieved by having no agreement.

Proposition 3 shows that agreements are not always good, and that one has to be careful when advocating particular agreements. Of course, at the emission stage, once the investments are sunk, all countries benefit from negotiating an agreement. It is the *anticipation* of negotiations which reduces investments and thus utility. Thus, if (4.10) holds, the countries would have been better off if they were committed to abstain from negotiating short-term agreements. A way of committing may be to agree on emissions in advance, before the investments occur.

4.3. Long-term Agreements

The hold-up problem under short-term agreements arose because the g_is were negotiated after investments were made. If the time horizon of an agreement is longer-lasting, however, it is possible for countries to develop technologies within the time-frame of the agreement. The other countries are then unable to hold up the investing country, since the quotas have already been negotiated (at least for the near future). For real-world long-term agreements, it is indeed reasonable that future commitments are made for such a long time horizon that a country is able to invest between the time at which the promises were made, and the time at which the last promise is supposed to be kept.

4.3.1. One-period Agreements

This interpretation of "long-term agreements" can be captured easily in the model: Just let the countries negotiate the g_is in the beginning of each period, before the investments are made. While these agreements only last one period, they are indeed "longer" than the short-term agreements studied above. Moreover, each period can be quite long in the model, since I have not specified whether the discount factor, for example, is large or small (multi-period agreements are nevertheless analyzed below).

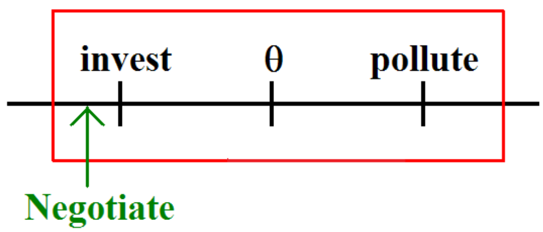

Figure 3: The timing for "long-term agreements"

Naturally, 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

$$B'\left(g_i^{lt} + R_i^{lt}\right) = \frac{K\left(1 - \delta q_R/n\right)}{D - e\left(n-1\right)}. \tag{4.11}$$

Equilibrium R_i^{lt} is thus large for small g_i^{lt}. Compared to (4.9), (4.11) suggests that countries invest more under long- than short-term agreements (at least for the same g_i). But compared to (4.2), countries still under-invests if $e > 0$ or $n > 1$. First, a country does not internalize the spillover e on the other countries. Second, if the agreement is not lasting forever ($\delta > 0$), a country anticipates that a good technology worsens its bargaining position in the future, once a new agreement is to be negotiated. A good technology, at that stage, 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.[27] 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 with 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 (4.11) implies that the equilibrium R_i^{lt}s are less than optimal, the g_i^{lt}s are sub-optimally low *ex post*. Combining (4.3) and (4.11), we can write

$$B' - \mathrm{E}nC' - n\delta U_G = \frac{K}{D}\left(\frac{e(1-\delta q_R)(n-1) + \delta q_R (1-1/n)}{D - e(n-1)}\right) \Rightarrow \tag{4.12}$$
$$g_i^{lt} = \mathrm{E}g_i^* - \frac{K}{D(b+cn^2)}\left(\frac{e(1-\delta q_R)(n-1) + \delta q_R (1-1/n)}{D - e(n-1)}\right) \text{ if (Q).}$$

Taking the investments as given, the g_i^{lt} should optimally have satisfied $B' - \mathrm{E}nC' - n\delta U_G = 0$ rather than (4.12). Only that would equalize marginal costs and benefits of abatement.[28] Relative to this *ex post* optimal level, g_i^{lt} satisfying (4.12) must be *lower*, particularly if the right-hand side of (4.12) is large, i.e., if e and δ are large. This makes the long-term agreement more demanding or *tougher* to satisfy at the emission stage. The purpose of such an over-ambitious agreement is to encourage investments, since these are sub-optimally low when e and δ are large.

Proposition 4. *(i) A country invests more if the agreement is tough (4.11). (ii) The optimal agreement (4.12) is therefore tougher if the externality e is large and the time horizon short (δ large). (iii) This constitutes the unique MPE, in which (4.7) holds.*

On the other hand, if $e = \delta q_R = 0$, the right-hand side of (4.12) is 0, meaning that the commitments under the 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 their technologies affect future bargaining power. Thus, investments are first best and there is no need to distort the g_i^{lt}s downwards.

[27] 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.

[28] As recommended e.g. by the Stern Review (Stern, 2007)

With this type of long-term agreement, $U_{R\_}$ and $U_{G\_}$ are exactly as in the two previous subsections. Thus, the predicted contract and investments are robust whether there is a long-term agreement, a short-term agreement, or no agreement in the subsequent period.

4.3.2. Multi-Period Agreements

Although the "long-term" agreement above lasts only a single period, that period can be quite long (if e.g. δ is small). Thus, it may not be a surprise that the intuition survives if agreements can last several periods. 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 (4.11) holds. Anticipating this, i can invest less in period T by investing more in period $T-1$, and the net investment cost is then $K(1-\delta q_R)$. The same logic applies to every previous period and, in equilibrium,

$$\mathrm{E}B'(g_{i,t} + R_{i,t}) = \frac{K(1-\delta q_R)}{d} = \frac{K(1-\delta q_R)}{D - e(n-1)} \text{ for } t < T. \tag{4.13}$$

Thus, the incentives to invest are larger earlier than in the last period (4.11). 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 $T+1$.[29]

This is all 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 $t \leq T$. Thus, the pollution level is similar to the first best, despite the suboptimally low investments. The $g_{i,t}$s are thus lower than what is optimal *ex post*. Combining (4.3) and (4.13) for $t < T$,

$$\begin{aligned} B' - \mathrm{E}nC' - n\delta U_G &= \frac{K}{D}\left(\frac{e(n-1)(1-\delta q_R)}{D-e(n-1)}\right) \Rightarrow \\ g_{i,t} &= \mathrm{E}g_i^* - \frac{K}{D(b+cn^2)}\left(\frac{e(n-1)(1-\delta q_R)}{D-e(n-1)}\right) \text{ if (Q).} \end{aligned} \tag{4.14}$$

Ex post, $B' - nC' - n\delta U_G = 0$ is optimal. Compared to this, $g_{i,t}$ satisfying (4.14) should be smaller if e is positive and large. This makes the agreement more demanding and tougher to satisfy ex post, inducing countries to invest more. For $t = T$, however, (4.12)

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

continues to hold and since its right-hand side is less than that of (4.14), $y_{i,T} < y_{i,t}$ for $t < T$. In words: the agreement should be tougher to satisfy towards the end in order to encourage investments.[30]

Proposition 5. *(i) For a T-period agreement, investments decrease towards the end. (ii) To encourage more investments, the agreement becomes tougher over time compared to the ex post optimum (4.12)-(4.14). (iii) This constitutes the unique MPE, in which (4.7) holds.*

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 if T is large.

Unfortunately, it is not possible to derive the optimal T in the general case. If B and C are quadratic (Q), however, Appendix shows:

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 is large. Reinforcing this problem by a small T is then especially harmful, and the optimal T is larger. The hold-up problem (and under-investment) is also greater if n is large and b small.[31] Naturally, T should be smaller if future optimal emissions are uncertain (σ large) and important (c large).

[30] Alternatively, countries could negotiate just before the emission stage, after θ_t, $t = 1$, is realized. The first-period emission levels would then be ex post optimal (as under short-term agreements) while later they would be lower and given by (4.3). Such an agreement should thus become tougher over time to encourage investments. If $T = 2$, this is equivalent to a short-term agreement in the first period and a "long-term" agreement for the second, exactly as they are analyzed in Sections 4.2 and 4.3.1.

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

4.4. Long-term Agreements with Renegotiation

The long-term agreements above are sub-optimal ex post. Not only are the commitments made before one knows the severity of the problem (determined by θ). In addition, they specify emission levels that are less than what is *expected* optimal ex post. The countries may thus be tempted to renegotiate the treaty, after θ and the investments are realized. In other words, long-term agreements are not renegotiation-proof. Suppose, therefore, that renegotiation is costless and the countries cannot commit to abstain from renegotiation.

4.4.1. One-period Agreements and Renegotiation

The timing is now the following. First, the countries negotiate the initial commitments g_i^{de} (called the "default outcome"). If these negotiations fail, it is natural to assume that the threat point is no agreement (just as before).[32] Thereafter, the countries invest and θ is realized. Before carrying out their commitments, the countries get together and renegotiate the g_i^{de}s. The bargaining surplus is split equally.

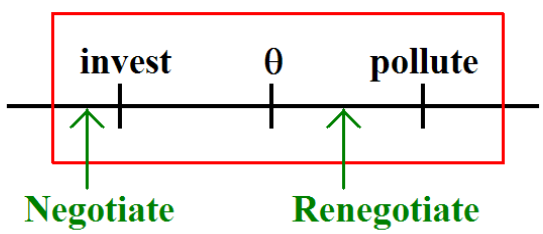

Figure 4: The timing when renegotiation is possible

What is the threat point if renegotiation fails? One possibility is that the countries get upset and revert to no cooperation. If so, the renegotiation game is identical to the negotiations under short-term agreements, and so are the incentives to invest. As argued above, this is worse than no agreement at all.

A more reasonable threat point, however, is the initial agreement. After all, when the Doha trade negotiations broke down, the countries reverted to the existing set of trade agreements, not the non-cooperative outcome. Thus, assume the g_i^{de}s are enforced unless renegotiation succeeds.

[32] If the threat point where short term agreements, negotiated after the investment stage, the outcome would be identical.

Renegotiation ensures that emission levels are ex post optimal, and not less as recommended in Section 4.3. When investing, a country anticipates that it does not, in the end, have to comply to an over-ambitious long-term agreement. Will this jeopardize the incentives to invest?

Proposition 7. *(i) With renegotiation, all investments and emissions are first best if just the initial agreement satisfies (4.15). The initial quotas (g_i^{de}) should thus be smaller if the spillover (e) is large and the time horizon short (δ large). (ii) This constitutes the unique MPE, in which (4.7) holds.*

$$B'\left(g_i^{de}+R_i^*\right) = \frac{K}{D-en} \Rightarrow \tag{4.15}$$

$$B'\left(g_i^{de}+R_i^*\right) = \mathrm{E}B'\left(g_i^*+R_i^*\right)+\frac{K}{D}\left(\frac{en}{D-en}+\delta q_R\right) \Rightarrow \tag{4.16}$$

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

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 happy with the (non-cooperative) *default* outcome in which they could pollute more. This made the "technology-losers" reluctant to negotiate, giving them a better bargaining position. Things are quite different when *re*negotiating an *ambitious* agreement. Then, the technology-losers are desperate to reach a new agreement, replacing the very expensive commitments. Such countries now have a poor bargaining position, and they are, in equilibrium, going to get a quite 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, particularly in situations where the countries otherwise are tempted to under-invest, i.e., if the externality e is large, and if δq_R is large, since then the countries fear that more technology today hurts their bargaining position in the near future. Thus, the agreement should be more ambitious if e and δq_R are large. Formally, (4.15) implies that g_i^{de} decreases in e and δ since R_i^* is increasing in δ but independent of e. No contract is helpful if $e \geq d$.[33]

Compared to (4.12), 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) trades off the concern for investments (by reducing g_i^{lt}) and the ex post optimum (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 - and so tough that investments are first best.

4.4.2. Multiple Periods and Renegotiation

The first best is implemented for any long-term agreements lasting $T \geq 1$ periods if renegotiation is possible. If $T = 1$ or at $t = T \geq 1$, the initial agreement should be as discussed above. If $T > 1$ and the agreement specifies $g_{i,t}^{de}$, $t \in \{1, ..., T\}$, then investments at $t < T$ is first best if just:

$$B'\left(g_{i,t}^{de} + R_i^*\right) = \frac{K\left(1 - \delta q_R\right)}{D - en} \Rightarrow g_{i,t}^{de} = \mathrm{E}g_{i,t}^* - \frac{K}{bD}\left(\frac{1 - \delta q_R}{D/en - 1}\right) \text{ if (Q).} \qquad (4.17)$$

Compared to (4.15), it is clear that g_i^{de} is larger when $T > 1$ than when $T = 1$ (R_i^* is the same in both cases). Thus, agreements lasting one period should be more ambitious than if $T > 1$, confirming the earlier observation that an agreement should be more ambitious if it is short-lasting. The agreement should also become more ambitious over time, since $g_{i,T}^{de}$ is given by (4.15).

Proposition 8. *(i) With renegotiation, the outcome is first best for any $T \geq 1$. If $T > 1$, the $g_{i,t}^{de}$s should be larger and given by (4.17) for $t < T$, thus still decreasing in e. (ii) $g_{i,t'}^{de}$, $t' > t + 1$, is irrelevant. (iii) This constitutes the unique MPE, in which (4.7) holds.*

Note that the first best is implemented at t no matter $g_{i,t'}^{de}$, $t' > t$. It is only the subsequent default commitments that matter for the incentives to invest. But since the optimal R_i^* depends on the last realized θ, the $g_{i,t+1}^{de}$s must be renegotiated after θ is realized in period t. A possible time is just after it is realized, once the countries also renegotiate the

[33] If $e \geq d$, $D \leq en$ and (4.15) implies that $B' \leq 0$, which may be impossible. Contracts are then not helpful (in line with Che and Hausch, 1999), since i's investment is then improving j's bargaining power rather than i's.

$g_{i,t}^{de}$s and replace them with the emission levels that will actually be undertaken. With this procedure, each time the countries negotiate, they are agreeing to the $g_{i,t+1}^{de}$s for the following period and these targets are more ambitious than what is expected to be optimal ex post. At the very same time, the countries are renegotiating their earlier ambitious promises, replacing them with higher emissions. This procedure is reminiscent of familiar time-inconsistency problems. But, in contrast to those stories, this procedure implements the first best.

Corollary 1. *At every t, the countries are (re)negotiating an ambitious agreement for $t+1$ while, at the same time, replacing the current agreement with higher, ex post optimal quotas.*

5. Trade, Tariffs and Intellectual Property Rights

This section introduces a new externality, relates it to trade agreements and analyzes the effects of and for climate treaties.

Externalities: There are several ways in which spillovers could be formalized, but many of them give similar results to those above.[34] An alternative (or addition) to the spillover above arises by assuming that j benefits *directly* by the externality xr_i when $i \neq j$ invests. If K continues to be the *social* net marginal investment cost, i's private investment cost is

$$k \equiv K + (n-1)\,x. \tag{5.1}$$

The model is unchanged if we just write the utility as

$$u_i = B(y_i) - C(G) - kr_i + \sum_{j \neq i} xr_j.$$

The externality x can be interpreted as a *general* technological spillover (affecting u_i and not only i's environmental technology) or as a spillover that reduces i's *cost* of making

[34] 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.

a particular investment r_i (such that the cost is $kr_i - \sum xr_j$).[35] The Appendix allows for both e and x and finds them to play similar roles.

Trade in technologies: The externality x may reflect international law. Suppose that r_i has the *potential* of reducing j's cost (or increasing u_j) by $\overline{x}r_i$ 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. If i sets the price, i charges j's willingness to pay, $(1 - \gamma)\, \overline{x}$, for each invested unit. Clearly, the *net* externality for j is $x = \gamma\overline{x}$. Weaker intellectual property rights mean larger γ, x and k.

Firms: Private firms have been ignored so far in the analysis. But since firms may develop most of the technology in reality, it is comforting to note that the results would not necessarily change if firms were introduced. If the government can perfectly regulate the firms' investments in technologies by specifying conditional fees or grants, then firms are perfect agents for the government and it is sufficient to consider the government's incentives.[36] Even without regulation, if the governments are outsourcing the development of technology and firms compete by setting prices, anticipating the revenues $(n - 1)(1 - \gamma)\, \overline{x}$ when licensing to foreigners, technological units are provided at cost-price k and firms are not affecting the game.

Tariffs and subsidies: If the foreign individuals or firms paying for the externality $(1 - \gamma)\, \overline{x}$ face an ad valorem tariff τ, they are willing to pay only $(1 - \tau)(1 - \gamma)\, \overline{x}$ for each imported unit. On the other hand, since trade is verifiable, one may consider encouraging R&D by subsidizing trade in abatement technology.[37] Let s represent this subsidy, paid for by either the importing country or uniformly by the non-exporting countries. In either case, the net externality for country j when country i decides to invest becomes

$$x = \gamma\overline{x} - (s - \tau)(1 - \gamma)\, \overline{x}.$$

The private cost of investing faced by a country (or government) is k given by (5.1), just

[35] 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.

[36] The necessary regulation may certainly not be trivial if e.g. researchers need to relocate between dirty and clean sectors. A combination of temporary input taxes, profit taxes, resource taxes and subsidies may then be necessary (Acemoglu *et al.* 2009).

[37] Stern (2007, p. 398) states "There are two types of policy response to spillovers... enforcement of private property rights through patenting [and] government funding."

as before. Thus, k and x are both higher with tariffs, since the importing country is then capturing more of the surplus, but lower if importers subsidize technological trade. The role of s would be identical if directed to investments rather than the associated trade.

Proposition 9. *If the subsidy (s) is low, tariff (τ) high and the intellectual property rights weak (γ large), (i) the agreement should be tougher and more long-lasting while (ii) short-term agreements are likely to be worse than no agreement under (Q).*

While the proof is in the Appendix, the intuition is straightforward. With tariffs, small subsidies and weak property right protection, firms do not capture the benefit experienced by the foreigners. This forces the government to pay more to compensate the firms when investing, and they invest less. A further reduction in investment is then particularly bad, making short-term agreements worse than the non-cooperative equilibrium (under (Q)). To encourage more investments, it is better to negotiate an agreement that is tougher and more long-lasting.

If s, τ or γ can be specified by international law, one may ask what their levels should be. Does the optimal subsidy, tariff and intellectual property right protection depend on the climate treaty?

Proposition 10. *s should be larger while τ and γ smaller if the agreement is short-lasting. The optimal s, τ and γ are given by (5.2) for short-term agreements, (5.3) for long-term agreements (and the last period of multi-period agreements), and by (5.4) for multi-period agreements (except for the last period).*

$$\left(s^{st}-\tau^{st}\right)\left(1-\gamma^{st}\right)-\gamma^{st} = \frac{K}{n\overline{x}} > \tag{5.2}$$

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

$$\left(s^{t}-\tau^{t}\right)\left(1-\gamma^{t}\right)-\gamma^{t} = \frac{Ke}{\overline{x}D}. \tag{5.4}$$

It is more important to encourage investments by protecting intellectual property rights, subsidizing technological trade and reducing tariffs if the climate treaty is short-lasting, since the hold-up problem is then larger. Such "trade agreements" are thus strategic substitutes to climate treaties: Weakening cooperation on one area makes further cooperation on the other more important.

If the subsidy can be freely chosen, short-term agreements are first best: while (5.2) induces optimal investments, countries are negotiating the ex post optimal emissions. The emission levels under long-term agreements (without renegotiation) are never first best, however, due to the stochastic θ.

Corollary 2. *If s can be freely chosen, short-term agreements are first-best while long-term agreements (without renegotiation) are not.*

If renegotiation is possible, the first best is feasible no matter s, τ and γ as long as the initial agreement is more ambitious for large e and x (and thus large γ and τ but small s). Under (Q), the g_i^{de}s should satisfy[38]

$$g_i^{de} = \mathrm{E}g_i^* - \frac{K}{bD}\left(\frac{x/K + e/D}{1/n - e/D} + \delta q_R\right). \tag{5.5}$$

If the g_i^{de}s are exogenously given and high (e.g. because a tough climate treaty would be impossible to enforce), efficiency and (5.5) are still fulfilled if just x is sufficiently small (requiring a large s and small τ or γ). This suggests that less ambitious climate treaties should be accompanied by technological subsidies, low tariffs and property right protection, confirming that the two types of agreements are strategic substitutes.

Corollary 3. *With renegotiation, the first best is implemented even if g_i^{de} increases, if just s increases or τ or γ decreases.*

6. Generalizations

6.1. Political Instruments

To help intuition and simplify the argument, above I have assumed that quotas cannot be traded and side payments are available. The proofs, however, do not rely on these assumptions.

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

[38] The general conditions are derived in the Appendix.

$B'(g_i + R_i)$ is the value of being allowed to pollute one more unit, keeping G and R constant. Under short-term agreements, since the bargaining outcome is always $y_i = y_j$, there cannot be any trade in permits and countries invest the same whether B' reflects the value of consuming energy or selling a superflous permit. Under long-term agreements, if all countries invest the same, $y_i = y_j$, there will be no trade in permits, and B' captures the marginal value of investing no matter what it represents. But if a country considers deviating by investing substantially more (less) than the equilibrium amount, the marginal value of polluting will decrease (increase) less when permits can be sold (bought) than when they cannot. In either case, however, the value decreases (increases) when i invests more (less) due to the impact on the equilibrium permit price. This reduces (increases) the incentive to invest, and a deviation is thus not desirable. In short, trade does not induce countries to increase (or reduce) investments compared to the equilibria above.[39]

Negotiating the initial quota allocation is just like negotiating side payments, so then it does not matter whether transfers are available in addition.

An earlier version of this paper analyzed emission taxes in addition to quotas. The results were similar. For example, the first best is feasible with renegotiation if just the initial tax is more ambitious that what is expected to be optimal ex post, particularly if the externality is large and the agreement short-lasting.

6.2. Heterogeneity

To simplify, I started out by assuming the countries were completely symmetric and there were 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. Thus, it is not a necessary assumption that all countries start out with the same technology. If these stocks differ, nothing is changed in the analysis above.

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 strategies and the bargaining game would remain the same. The realized

[39] 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.

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 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 point (not necessarily bliss). Recognizing the importance of such heterogeneity, the Appendix does allow the reference point $\overline{y}_i$ to vary. While a large $\overline{y}_i$ increases the equilibrium g_i, the comparative static is unchanged.

Proposition 12. *All results continue to hold if countries have different bliss or reference point $\overline{y}_i$ for energy consumption.*

Other forms for heterogeneity are 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 the investing countries would as before invest too little. In a long-term agreement, one could encourage these countries to invest more by reducing their g_i^{lt} or, if renegotiation is possible, g_i^{de}. Such small g_is would not be necessary (nor optimal) for non-investing 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 non-investing countries with higher g_js. Thus, with heterogeneity in investment costs, it matter a great deal whether side transfers are possible and permits tradable. Proposition 11 is false if such heterogeneity is introduced. Evaluating political instruments under heterogeneity is thus important for future research.[40,41]

[40] For example, Buonanno *et al.* (2003) finds that with heterogeneity and tradable permits, some countries may over-invest to affect the equilibrium price for permits.

[41] Heterogeneities in the externality, country-size and the cost of climate change are also immensely important to investigate in future research, particularly when the externality follow from country-specific trade policies.

7. 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, I find a tractable and unique Markov Perfect Equilibrium (MPE) despite the $1+n$ stocks. While the model has several applications, the implications for climate treaties are particularly important.

The larger is one country's stock of abatement technology, the more the other countries choose to pollute and the less they invest. The more one country pollutes, the less the other countries pollute in the future, and the more they will invest. Both effects induce countries to pollute more and invest less than they would in a related static model. Since the unique MPE does not permit self-enforcing agreements, I consider various possibilities to negotiate and commit.

While the non-cooperative outcome is bad, a sequence of short-term agreements can be worse. At the negotiation stage, a country with good technology is going to be hold up by the others, requiring it to reduce its pollution by a lot. Anticipating this, countries invest less when negotiations are anticipated. The countries become worse off relative to the situation with no agreement, particularly if the number of countries is large.

The optimal long-term agreement is ambitious and specifies emission levels that are less than what is ex post optimal, particularly if it is relatively short-lasting and the technological spillovers large. The length of an agreement, if it can be freely chosen, should be longer if the spillover is large.

But such long-term agreements are not renegotiation-proof. If renegotiation is possible and cannot be prevented, the outcome is not worse but first best. Renegotiation ensures that the emission levels are ex post optimal, and the role of the initial agreement is only to affect the incentives to invest. Again, the agreement should be more ambitious if its time horizon is short and the externalities large. Flexibility regarding future commitments is thus best ensured by renegotiating ambitious long-term agreements rather than by letting them quickly expire.

The results hold no matter whether side transfers are available, permits tradable, firms included or the technology can be traded, taxed or subsidized. If trade in abatement technology face high tariffs or low subsidies, the climate treaty should be tougher and last longer. If the climate treaty is short-lasting and not ambitious, investments or trade in technology should be subsidized, tariffs reduced and intellectual property rights strengthened. Trade and climate treaties are thus strategic substitutes.

While this paper establishes some benchmark results, it is only a small step towards a better understanding of good environmental agreements. I have not distinguished between innovation and diffusion, and I have abstracted from domestic politics, heterogeneity, private information, monitoring, compliance and the possibility to opt out of the agreement. Relaxing these assumptions are the natural next steps.

8. Appendix

All propositions are here proven with the generalizations discussed in Sections 5-6: 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 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, $B(y_i) \equiv \beta(y_i - \overline{y}_i)$ 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 $m \equiv -\delta\partial U_i/\partial G_-$, $z = \delta\partial U_i/\partial R_-$, $\widetilde{R} \equiv q_R R_-$, $\widetilde{G} \equiv q_G G_- + \theta$, 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 $\overline{y}_i$s, and the payoff relevant states are G and R. Analyzing the symmetric equilibrium (where symmetric countries invest the same amount), I thus 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 proof for Proposition 1.

8.1. Proposition 1 (no cooperation)

At the emission stage, each country's first-order condition is (when choosing y_i):

$$\begin{aligned} 0 &= \beta'(y_i - \overline{y}_i) - C'(G) + \delta U_G(G, R) \\ &= \beta'(\widetilde{y}_i - \overline{y}) - C'\left(\widetilde{G} - R + \sum \widetilde{y}_i\right) + \delta U_G(\widetilde{G} - R + \sum \widetilde{y}_i, R), \end{aligned} \tag{8.1}$$

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

$$\mathrm{E}W(\widetilde{G}, R) - kr_i = \mathrm{E}W\left(q_G G_- + \theta, \widetilde{R} + \sum_i Dr_i\right) - kr_i, \tag{8.2}$$

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 $(R(G_-) - q_R R_-)/Dn$, and thus:

$$U(G_-, R_-) = \mathrm{E}W(q_G G_- + \theta, R(G_-)) - (k - (n-1)x)\left(\frac{R(G_-) - q_R R_-}{Dn}\right) \Rightarrow \tag{8.3}$$

$$z/\delta \equiv \frac{\partial U}{\partial R_-} = \frac{q_R K}{Dn} \tag{8.4}$$

in every period. Hence, $U_{RG} = U_{GR} = 0$, m and U_G cannot be functions of R and (8.1) 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. Thus, write $G\left(\widetilde{G} - R\right)$. (8.2) 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:

$$\begin{aligned} U(G_-, R_-) &= \mathrm{E}\gamma(\xi + \theta) - Kr + \delta U(G(\xi+\theta), R) \\ &= \mathrm{E}\gamma(\xi + \theta) - K\left(\frac{q_G G_- - \xi - q_R R_-}{Dn}\right) + \delta U(G(\xi+\theta), q_G G_- - \xi) \Rightarrow \\ m/\delta &= \partial U/\partial G_- = -K\left(\frac{q_G}{Dn}\right) + \delta U_R q_G = -\frac{K q_G}{Dn}(1 - \delta q_R), \end{aligned} \tag{8.5}$$

since $G(\xi + \theta)$ and $\gamma(.)$ are *not* functions of G_- when $q_G G_- - R = \xi$. Since U_G is a constant, (8.1) implies that if R increases, $\widetilde{y}_i$ increases but G must decrease, implying $\partial\widetilde{y}_i/\partial R \in (0,1)$, 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 $i = j$.

Under (Q), (8.1) becomes

$$0 = -cG + b\overline{y}_i - by_i - m \Rightarrow y_i = \overline{y}_i - \frac{m + cG}{b}. \tag{8.6}$$

$$G = \sum_j (y_j - R_j) + \widetilde{G} = \widetilde{G} + n\left(\overline{y} - \frac{m + cG}{b}\right) - \sum_j R_j \Rightarrow$$

$$G = \frac{b\overline{y}n - mn + b\left(\widetilde{G} - R\right)}{b + cn}, \text{ so} \tag{8.7}$$

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

$$g_i = y_i - R_i = (\overline{y}_i - \overline{y}) + \frac{b\overline{y} - m - c\left(\widetilde{G} - \sum_{j \neq i} R_j\right)}{b + cn} - \frac{R_i(b + cn - c)}{b + cn}.$$

Interrim utility (after investments are sunk) can be written as:

$$\begin{aligned} W_i^{no} &\equiv -c\left(1+c/b\right)G^2/2 - Gmc/b + \frac{\left(b\overline{y}\right)^2 - m^2}{2b} + \delta U(G,R). \text{ Thus,} \\ \partial W_i^{no}/\partial R &= c\left(1+c/b\right)G\left(\frac{b}{b+cn}\right) + \frac{bm\left(1+c/b\right)}{b+cn} + z. \qquad (8.8) \end{aligned}$$

At the investment stage, each country sets $k/D = \mathrm{E}\partial W_i^{no}/\partial R$. From (8.8):

$$\begin{aligned} k/D &= \mathrm{E}c\left(G\right)\left(\frac{b+c}{b+cn}\right) + \frac{m\left(b+c\right)}{b+cn} + z \Rightarrow \\ \mathrm{E}G &= \frac{k\left(b+cn\right)}{cD\left(b+c\right)} - \frac{m}{c} - \frac{z}{c}\left(\frac{b+cn}{b+c}\right), \text{ combined with (8.7):} \qquad (8.9) \\ R &= q_G G_- + n\overline{y} - nm/b - \frac{b+cn}{b}\mathrm{E}G \\ &= q_G G_- - \frac{k\left(b+cn\right)^2}{Dcb\left(b+c\right)} + \overline{y}n + \frac{z\left(b+cn\right)^2}{cb\left(b+c\right)} + \frac{m}{c} \Rightarrow \qquad (8.10) \\ r_i n D &= -q_R R + q_G G_- - \frac{k\left(b+cn\right)^2}{Dcb\left(b+c\right)} + \overline{y}n + \frac{z\left(b+cn\right)^2}{cb\left(b+c\right)} + \frac{m}{c}. \end{aligned}$$

Since $\widetilde{G} = q_G G_- + \theta$, (8.7) gives $G = \mathrm{E}G + \theta b/\left(b+cn\right)$. Substituting in (8.6) and (8.9):

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

This is helpful when calculating u_i^{no}:

$$\begin{aligned} u_i^{no} &= -\frac{c}{2}\left(\frac{k\left(b+cn\right)}{Dc\left(b+c\right)} - \frac{m}{c} - \frac{z\left(b+cn\right)}{c\left(b+c\right)} + \frac{\theta b}{b+cn}\right)^2 - \frac{b}{2}\left(\frac{\left(k/D - z\right)\left(b+cn\right)}{b\left(b+c\right)} + \frac{\theta bc}{b\left(b+cn\right)}\right)^2 \\ &\quad - \frac{K}{Dn}\left(-\widetilde{R} + q_G G_- - \frac{k\left(b+cn\right)^2}{Dcb\left(b+c\right)} + \overline{y}n + \frac{z\left(b+cn\right)^2}{cb\left(b+c\right)} + \frac{m}{c}\right) \Rightarrow \\ \mathrm{E}u_i^{no} &= -\frac{1}{2}\left(\frac{k}{D} - z\right)^2\left(\frac{b+cn}{b+c}\right)^2\left(\frac{1}{c} + \frac{1}{b}\right) - \frac{m^2}{2c} + \frac{m}{c}\left(\frac{b+cn}{b+c}\right)\left(\frac{k}{D} - z\right) \\ &\quad - \frac{K}{Dn}\left(q_G G_- - \widetilde{R} - \frac{\left(b+cn\right)^2}{bc\left(b+c\right)}\left(\frac{k}{D} - z\right) + \overline{y}n + \frac{m}{c}\right) - \frac{bc\left(b+c\right)\sigma^2}{2\left(b+cn\right)^2}. \end{aligned}$$

8.2. Proposition 2 (short-term agreements)

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 (8.1) requires:

$$0 = \beta'\left(\widetilde{y}_i - \overline{y}\right)/n - C'\left(\widetilde{G} - R + \sum \widetilde{y}_i\right) + \delta U_G(\widetilde{G} - R + \sum \widetilde{y}_i, R). \qquad (8.11)$$

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:

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

Combined with (8.11),

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

Under (Q), the first-order conditions for the optimal $\widetilde{y}_i$s:

$$\begin{aligned}
0 &= -ncG + b\overline{y} - b\widetilde{y}_i - nm \Rightarrow \widetilde{y}_i = \overline{y} - \frac{nm+ncG}{b}. \\
G &= \sum_j (y_j - R_j) + \widetilde{G} = \widetilde{G} + n\left(\overline{y} - \frac{nm+ncG}{b}\right) - R \Rightarrow && (8.12) \\
G &= \frac{b\overline{y}n - mn^2 + b\left(\widetilde{G} - R\right)}{b+cn^2}, \text{ so} && (8.13) \\
y_i &= \overline{y}_i - \overline{y} + \frac{b\overline{y} - mn - cn\left(\widetilde{G} - R\right)}{b+cn^2} \text{ and} \\
g_i &= \overline{y}_i - \overline{y} + \frac{b\overline{y} - mn - cn\left(\widetilde{G} - R\right)}{b+cn^2} - R_i. && (8.14)
\end{aligned}$$

Interrim utility is

$$\begin{aligned}
W_i^{st} &= -\frac{c}{2}G^2 - \frac{b}{2}\left(\frac{nm+ncG}{b}\right)^2 + \delta U(G,R), \text{ so} \\
\partial W_i^{st}/\partial R &= cG + m + z.
\end{aligned}$$

A country invests until the marginal costs of investment is

$$k = D(\mathrm{E}cG + m + z) \Rightarrow \mathrm{E}G = \frac{k}{cD} - \frac{m+z}{c}. \qquad (8.15)$$

Subsituting in (8.12), after taking the expection of it, and solving for R gives

$$R = q_G G_- + n\overline{y} + \frac{m}{c} - \left(\frac{b+cn^2}{b}\right)\left(\frac{k}{cD} - \frac{z}{c}\right). \tag{8.16}$$

From (8.13) and (8.15):

$$\begin{aligned}
G &= \frac{k}{cD} - \frac{m+z}{c} + \frac{b\theta}{b+cn^2} \Rightarrow \qquad (8.17)\\
\overline{y} - \widetilde{y}_i &= \frac{nm}{b} + \frac{nc}{b}\left(\frac{k}{cD} - \frac{m+z}{c} + \frac{b\theta}{b+cn^2}\right) = \frac{n}{b}\left(\frac{k}{D} - z + \frac{bc\theta}{b+cn^2}\right) \Rightarrow\\
u_i^{st} &= -\frac{c}{2}G^2 - \frac{b}{2}\left(\overline{y} - \widetilde{y}_i\right)^2 - Kr\\
&= -\frac{c}{2}\left(\frac{k}{cD} - \frac{m+z}{c} + \frac{\theta b}{b+cn^2}\right)^2 - \frac{n^2}{2b}\left(\frac{k}{D} - z + \frac{\theta bc}{b+cn^2}\right)^2 - Kr \Rightarrow\\
\mathrm{E}u_i^{st} &= -\frac{1}{2}\left(\frac{k}{D} - z\right)^2\left(\frac{1}{c} + \frac{n^2}{b}\right) - \frac{m^2}{2c} + \frac{m\left(k/D - z\right)}{c}\\
&\quad -\frac{K}{nD}\left(q_G G_- - q_R R_- + n\overline{y} + \frac{m}{c} - \left(\frac{b+cn^2}{b}\right)\left(\frac{k}{cD} - \frac{z}{c}\right)\right) - \frac{\sigma^2 bc}{2\left(b+cn^2\right)}.
\end{aligned}$$

8.3. Proposition 3 (a comparison)

Comparing (8.10) with (8.16) and (8.9) with (8.17),

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

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

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

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

$$\mathrm{E}u_i^{st} - \mathrm{E}u_i^{no} + m\frac{(k/D - z)(n-1)}{(b+c)} - z\frac{k(n-1)^2}{D(b+c)}\left(1 - \frac{\delta q_R K}{nk}\right) =$$

$$\left(\frac{(bc\sigma)^2[n-1]^2}{2(b+nc)^2(b+cn^2)} - \frac{[n-1]^2}{2(b+c)}\left[\left(\frac{k}{D} - z\right)^2 - \frac{2K}{Dn}\left(\frac{k}{D} - z\right) + \frac{2zk}{D}\left(1 - \frac{\delta q_R K}{nk}\right)\right]\right) > 0$$

$$\Rightarrow \frac{(bc\sigma)^2(b+c)}{(b+nc)^2(b+cn^2)} > \left(\frac{K}{D}\right)^2\left[\left(\frac{k}{K}\right)^2 + \left(\frac{\delta q_R}{n}\right)^2 - \frac{2k}{nK} + \frac{2\delta q_R}{n^2} - \frac{2(\delta q_R)^2}{n^2}\right]$$

$$\begin{aligned} &= \left(\frac{K}{D}\right)^2\left[\left(\frac{K+(n-1)x}{K}\right)^2 - \left(\frac{\delta q_R}{n}\right)^2 - \frac{2(K+(n-1)x)}{nK} + \frac{2\delta q_R}{n^2} \pm \frac{1}{n^2}\right] \\ &= \left(\frac{K}{Dn}\right)^2\left[(n-1)^2\left(1+\frac{nx}{K}\right)^2 - (1-\delta q_R)^2\right]. \end{aligned}$$

8.4. Proposition 4 (long-term agreements)

When g_i is already negotiated, i invests until

$$\begin{aligned} k &= \beta'(g_i + R_i - \overline{y}_i)d + zD \Rightarrow \qquad (8.18) \\ \widetilde{y}_i - \overline{y} &= \beta'^{-1}\left(\frac{k-zD}{d}\right),\ R_i = \beta'^{-1}\left(\frac{k-zD}{d}\right) + \overline{y}_i - g_i, \\ dr_i &= \beta'^{-1}\left(\frac{k-zD}{d}\right) + \overline{y}_i - g_i - q_R R_{i,-} - \sum_{j\neq i} er_j. \qquad (8.19) \end{aligned}$$

Anticipating this, the utility before investing is:

$$U_i = \beta\left(\beta'^{-1}\left(\frac{k-zD}{d}\right)\right) - \mathrm{E}C(G) - kr_i + \sum_{j\neq i} xr_j + \delta U(G,R). \qquad (8.20)$$

If the negotiations fail, the default outcome is the non-cooperative outcome, giving everyone the same utility. Since the r_is follow from the g_is in (8.19), 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. (8.19) 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 $\partial r_i/\partial\zeta = -1/D \forall i$):

$$\begin{aligned} -n\mathrm{E}C'(G) + K/D + n\delta U_G - nD\delta U_R(1/D) &= 0 \Rightarrow \\ \mathrm{E}C'(G) + m + z &= K/Dn. \end{aligned} \tag{8.21}$$

Combined with (8.19), neither G nor R can be functions of R_- (R_i in (8.19) and (8.21) 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 $U_{RG} = 0$). (8.21) then implies EG is a constant and, since we must have $\zeta = (\mathrm{E}G - q_G G_-)/n + q_R R_-/n - \overline{y}$, (8.19) 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 and (8.4) and (8.5), just as before. Substituted in (8.21):

$$\mathrm{E}C'(G) = (1-\delta q_G)(1-\delta q_R)K/Dn. \tag{8.22}$$

This is the same pollution level as in the first best (4.3). But investments might be suboptimally low. Combining (8.22) with (8.18),

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

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

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

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

8.5. Proposition 5 (multiple periods)

At the start of $t = 1$, countries negotiate emission levels for every period $t \in \{1, ..., T\}$. The investment level in period T is (8.19) 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:

$$\begin{aligned} k &= d\beta'(g_{i,T-1} + R_{i,T-1} - \overline{y}_i) + \delta q_R k \Rightarrow \\ R_{i,T-1} &= q_R R_{i,T-1} + dr_{i,T-1} + \sum_{j\neq i} er_{j,T-1} = \overline{y}_i - g_{i,T-1} - \beta'^{-1}(k(1-\delta q_R)/d). \end{aligned} \tag{8.23}$$

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

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

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

$$\begin{aligned} Dr &= \left(\overline{y}_i - g_i - q_R\left(\overline{y}_i - g_i^{t-1} - \beta'^{-1}\left(k\left(1-\delta q_R\right)/d\right)\right)\right) - \beta'^{-1}\left(k\left(1-\delta q_R\right)/d\right) \\ &= \overline{y}_i\left(1-q_R\right) - g_i + q_R g_i^{t-1} - \left(1-q_R\right)\beta'^{-1}\left(k\left(1-\delta q_R\right)/d\right). \end{aligned} \tag{8.24}$$

Since $r_i = r_j$, (8.23) 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 (8.24), 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 (8.23),

$$\begin{aligned} \beta' - \mathrm{E}C'\left(G\right)n - nm &= \frac{k\left(1-\delta q_R\right)}{d} - \left(1-\delta q_R\right)\left(1-\delta q_G\right)K/D - \frac{\delta q_G\left(1-\delta q_R\right)K}{D} \\ = \frac{k\left(1-\delta q_R\right)}{d} - \left(1-\delta q_R\right)K/D &= \left(\frac{k}{d} - \frac{K}{D}\right)\left(1-\delta q_R\right) = \frac{K}{D}\left(\frac{x/K + e/D}{1/\left(n-1\right) - e/D}\right)\left(1-\delta q_R\right). \end{aligned}$$

The $g_{i,T}$ satisfies (8.22) 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), $\beta' = b\left(\overline{y}_i - g_i - R_i\right)$, $\mathrm{E}C' = c\left(\mathrm{E}\widetilde{G} + \sum g_j\right)$, and since $\beta' - \mathrm{E}cGn - n\delta U_G = 0$ for $g_i^*(R)$, we have

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

8.6. Proposition 6 (optimal length of agreement)

The optimal T trades off two costs: The first is associated with the hold-up problem, the second with the uncertain nature of the problem.

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):

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

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

The second 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,

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

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

$$
\begin{aligned}
L(T) &= \sum_{t=1}^{T} \frac{c}{2}\sigma^2 \delta^{t-1} \left(\frac{1-q_G^{2t}}{1-q_G^2} \right) = \frac{c\sigma^2}{2\left(1-q_G^2\right)} \sum_{t=1}^{T} \delta^{t-1}\left(1-q_G^{2t}\right) \\
&= \frac{c\sigma^2}{2\left(1-q_G^2\right)} \left[\frac{1-\delta^T}{1-\delta} - q_G^2 \left(\frac{1-\delta^T q_G^{2T}}{1-\delta q_G^2} \right) \right] \Rightarrow \qquad (8.25) \\
L'(T) &= \frac{c\sigma^2\left(-\delta^T \ln\delta\right)}{2\left(1-q_G^2\right)} \left[\frac{1}{1-\delta} - \frac{q_G^{2T+2}\left(1+\ln\left(q_G^2\right)/\ln\delta\right)}{1-\delta q_G^2} \right].
\end{aligned}
$$

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

$$
\begin{aligned}
&\min_T L(T) + \left(\delta^{T-1}H + \delta^T L\left(\widehat{T}\right)\right)\left(\sum_{\tau=0}^{\infty}\delta^{\tau\widehat{T}'}\right) \Rightarrow \\
&0 = L'(T) + \delta^T \ln\delta\left(H/\delta + L\left(\widehat{T}\right)\right) = L'(T) + \delta^T \ln\delta\left(H/\delta + L\left(\widehat{T}\right)\right) \\
= \quad & -\delta^T \ln\delta \left[\frac{c\sigma^2}{2\left(1-q_G^2\right)} \left(\frac{1}{1-\delta} - \frac{q_G^{2T+2}\left(1+\ln\left(q_G^2\right)/\ln\delta\right)}{1-\delta q_G^2} \right) - \frac{H/\delta + L\left(\widehat{T}\right)}{1-\delta^{\widehat{T}'}} \right] \qquad (8.26)
\end{aligned}
$$

assuming some T satisfies (8.26). 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 concave everywhere). Since the history (G_- and R_-) does not enter (8.26), T satisfying (8.26) is equal $\widehat{T}$, assuming also $\widehat{T}$ is optimal. Substituting $\widehat{T} = T$ and (8.25) in (8.26) gives:

$$
H/\delta = \frac{c\sigma^2 q_G^2}{2\left(1-q_G^2\right)\left(1-\delta q_G^2\right)} \left(\frac{1-\delta^T q_G^{2T}}{1-\delta^T} - q_G^{2T}\left(1 + \frac{\ln\left(q_G^2\right)}{\ln\delta}\right) \right), \qquad (8.27)
$$

increasing in T. $T = \infty$ is optimal if the left-hand side of (8.27) is larger than the right-hand side even if $T \to \infty$:

$$
\frac{c\sigma^2 q_G^2}{2\left(1-q_G^2\right)\left(1-\delta q_G^2\right)} < H/\delta. \qquad (8.28)
$$

If e (for given D), x (for given K) and n are large, but b small, H is large and (8.28) is more likely to hold and if it does not, the T satisfying (8.27) is larger. If c or σ^2 are large, (8.28) is less likely to hold and if it does not, (8.27) requires T to decrease.

8.7. Proposition 7 (long-term agreements with renegotiation)

Under 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 to its default utility, i's utility can be written as:

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

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.

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

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,

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

Combined with the optimum, (4.2),

$$\begin{aligned} \beta'\left(g_i^{de} + R_i^* - \overline{y}_i\right) - \mathrm{E}\beta'\left(g_i^* + R_i^* - \overline{y}_i\right) &= \frac{kn-K}{dn-D} - \frac{K}{D}(1-\delta q_R) \\ &= \frac{K}{D}\left(\frac{x/K + e/D}{1/n - e/D} + \delta q_R\right). \qquad (8.29) \end{aligned}$$

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 the same 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

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

8.8. Proposition 8 (multiple periods with renegotiation)

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,

$$\begin{aligned} k\left[1-\delta q_R\left(1-1/n\right)\right] &= (d-e)\left[\beta'\left(g_{i,1}^{de}+R_i^*-\overline{y}_i\right)\right](1-1/n)+K/n \\ \beta'\left(g_{i,1}^{de}+R_i^*-\overline{y}_i\right) &= \frac{k\left[n-\delta q_R\left(n-1\right)\right]-K}{(d-e)(n-1)}. \end{aligned} \tag{8.30}$$

Comparing (8.30) and (**??**) 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 $t-1$.

8.9. Proofs of Proposition 9-10 (optimal subsidises)

Proposition 9 is proven above since $x>0$ were allowed. 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

$$K+x\left(n-1\right)=K/N\Rightarrow -x=\frac{K}{n}.$$

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

$$\begin{aligned} K/D-nz &= k/d-zD/d=\left(K+x\left(n-1\right)\right)/d-zD/d\Rightarrow \\ -x &= \frac{K}{n}\left(\delta q_R+\frac{en}{D}\left(1-\delta q_R\right)\right). \end{aligned}$$

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

$$K\left(1-\delta q_R\right)/D=k\left(1-\delta q_R\right)/d\Rightarrow -x=eK/D.$$

The corollary follows immediately when recognizing that for short-term agreements both investments and emissions are first best at their respective stages.

8.10. Proofs of Propositions 11-12

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. The second part of Proposition 11 follows by recognizing that, when permits are tradable, altering their allocation is a form of side transfer, making the feasibility of explicit transfers irrelevant. Proposition 12 needs no further proof since heterogeneity is allowed in the proofs above.

References

Acemoglu, Daron; Aghion, Philippe; Bursztyn, Leonardo and Hemous, David (2009): "The Environment and Directed Technical Change," Mimeo, Harvard University.

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

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

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

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

Barrett, Scott and Stavins, Robert (2003): "Increasing Participation and Compliance in International Climate Change Agreements," *International Environmental Agreements: Politics, Law and Economics* 3: 349-376.

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

Battaglini, Marco and Coate, Stephen (2008): "A Dynamic Theory of Public Spending, Taxation, and Debt," *American Economic Review* 98 (1): 201-36.

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

Buonanno, Paulo; Carraro, Carlo and Galeotti, Marzio (2003): "Endogenous induced technical change and the costs of Kyoto," *Resource and Energy Economics* 25 (1): 11-34.

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.

Dockner, Engelbert J. and Long, Ngo Van (1993): "International Pollution Control: Cooperative versus Noncooperative Strategies," *Journal of Environmental Economics and Management* 24: 13-29.

Dockner, Engelbert J. and Sorger, Gerhard (1996): "Existence and Properties of Equilibria for a Dynamic Game on Productive Assets," *Journal of Economic Theory* 71: 209-27.

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.

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.

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 (2004): "Unilateral Emission Reductions and Cross-Country Technology Spillovers." *Advances in Economic Analysis & Policy* 4 (2), Article 3.

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.

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

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

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

Karp, Larry S. and Zhao, Jinhua (2008): "A Proposal for the Design of the Successor to the Kyoto Protocol." Mimeo, UC Berkeley (and winner of paper competition, Harvard Project on International Climate Agreements).

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

Leahy, Dermot and Neary, Peter (1997): "Public Policy Towards R&D in Oligopolistic Industries," *American Economic Review* 87 (4): 642-62.

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 (1987): "A Theory of Dynamic Oligopoly, III: Cournot Competition," *European Economic Review* 31: 947-68.

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

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

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

Nordhaus, William D. and Yang, Zili (1996): "A Regional Dynamic General-Equilibrium Model of Alternative Climate-Change Strategies," *American Economic Review* 86 (4): 741-86.

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 Zeeuw, Art de (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.

Rowat, Colin (2007): "Non-Linear Strategies in a Linear Quadratic Differential Game," *Journal of Economic Dynamics and Control* 31 (10): 3179-3202.

Rubio, Santiago J. and Ulph, Alistar (2007): "An infinite-horizon model of dynamic membership of international environmental agreements," *Journal of Environmental Economics and Management* 54 (3): 296-310.

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

Sorger, Gerhard (1998): "Markov-Perfect Nash Equilibria in a Class of Resource Games," *Economic Theory* 11: 79-100.

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

Spence, A. Michael (1977): "Entry, Capacity, Investment and Oligopolistic Pricing," *Bell Journal of Economics* 8: 534-44.

Stern, Nicholas (2007): "The Economics of Climate Change - The Stern Review", Cambridge University Press.

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

Wirl, Franz (1996): "Dynamic voluntary provision of public goods: Extension to nonlinear strategies", *European Journal of Political Economy* 12 (3): 555-60.

Wirl, Franz (2008): "Tragedy of the Commons in a Stochastic Game of a Stock Externality," *Journal of Public Economic Theory* 10 (1): 99-124.

Xepapadeas, Anastasios (1995): "Induced technical change and international agreements under greenhouse warming," *Resource and Energy Economics* 17 (1): 1-23.

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