Pledge and implement bargaining in the Paris Agreement on climate change

Alejandro Caparrós
Institute for Public Goods and Policies (IPP)
Consejo Superior de Investigaciones Científicas (CSIC)
Albasanz 26, 28037 Madrid, Spain
alejandro.caparros@csic.es

July, 2019

Abstract: This paper analyzes a multilateral bargain game motivated by the Paris Agreement on climate change. Each country submits pledges, which can be revoked but with reputational costs. Pledges detail intended abatement efforts and can be conditional or unconditional, depending on whether they depend on transfers. Pledges do not need to be accepted by other countries, although countries may negotiate agreements at the implementation phase. As this process is repeated over time, the paper analyzes the role of incomplete long-term provisions. The pledge and implement mechanism can implement the first-best in the short-term, and the surplus is shared according to the relative importance of the cost-of-revoking the pledges. However, without a long-term provision there is underinvestment. Even if the agreement does not cover investments, a long term provision contingent on the state of nature can bring the world to the long-term first-best, and a non-contingent provision can bring it closer.

JEL classifications: C73, C78, H41, Q54.
Key words: Bargaining, renegotiation, contract theory, climate change, international environmental agreements.

1 Introduction

Most countries of the world have signed and ratified the Paris Agreement. According to this agreement, countries aim at "holding the increase in the global average temperature to well below 2 °C above pre-industrial levels and [..] pursue efforts to limit the temperature increase to 1.5 °C above pre-industrial levels" (UNFCCC, 2015). The Paris Agreement does not have short-term commitments, but rather urges countries to send their voluntary commitments, their Nationally Determined Contributions (NDC). NDCs are not legally binding and are neither subject to the acceptance by other countries, nor to the unanimity rule needed to adopt international agreements, such as the Paris Agreement itself.

Rogelj et al. (2016) published an analysis determining the contribution of submitted Intended[1] NDCs (INDC) to the 2°C target (NDCs finally submitted were similar to the INDCs analyzed in this paper in most cases). The analysis shows that current climate policies cover about 25% of the required effort and that current INDCs would add another 25%, so that roughly 50% of the required effort is missing. In addition, about one third of the INDCs are conditional, requiring, for example, the provision of international funds[2], and an additional 45% of the INDCs came with both conditional and unconditional components. Thus, only about 20% of INDCs are purely unconditional (i.e. pledging to abate emissions by a certain amount irrespectively of the behavior of others). In other words, even if countries meet their (unconditional) INDCs there is a need for substantial additional effort to reach the 2 °C target (let alone the 1.5 °C target).

Although initial attention focused mainly on the role of the NDCs, the Paris Agreement has a second key feature: the mechanism to foster cooperation provided by articles 6 and 9, which may become the most important articles in the long term (Stavins and Stowe,

[1] INDCs were submitted before the Paris summit. Under the Paris Agreement, future mitigation contributions are referred to as NDCs, without the 'intended'.

[2] Although this practice was abandoned in the Agreement, as part of the Cancun agreement, a forerunner of the Agreement, the EU pledged to increase its effort "provided that other developed countries commit themselves to comparable emission reductions and that developing countries contribute adequately according to their responsibilities and respective capabilities".

2017). Article 6 allows countries to cooperate voluntarily to increase the ambition of their abatement targets. Cooperation between countries can significantly reduce costs and allow for more ambitious goals. Agreements based on this article can be legally binding and will probably have a shorter time horizon. Article 9 is also relevant as it considers monetary transfers from developed countries to assist developing countries in their efforts to mitigate and adapt to climate change. Developed countries have agreed to mobilize a substantial amount of money: "prior to 2025 the Conference of the Parties [. . .] shall set a new collective quantified goal from a floor of USD 100 billion a year".

The pledges system instituted in the Paris Agreement is usually referred to as the "pledge-and-review" mechanism. This terminology was used in the early nineties when debating about the appropriateness of moving from the "pledge-and-review" system introduced in the United Nation Framework Convention on Climate Change (UNFCCC) to the burden-sharing approach instituted in the Kyoto Protocol (a protocol to the UNFCCC). However, the system introduced in the Paris Agreement (another development of the UNFCCC) goes beyond a mere pledge-and-review mechanism. In addition to articles 6 and 9 mentioned above, at COP24 in Katowice national negotiators were charged with developing the so-called 'rulebook' to fully shape and specify the process set out by the Paris Agreement. The so-called 'Katowice package' is so detailed that some have argued that is has brought the Paris Agreement framework closer to a Kyoto-like framework. I therefore refer to the mechanism created by the Paris Agreement as "pledge-and-implement", in the sense that, after the submission of the pledges, additional negotiations should be expected during its implementation, mainly as developments of articles 6 and 9.

For illustration, let us focus on a developing country that requested in its conditional pledge a transfer in exchange of an additional reduction in its GHG emissions, beyond what it has pledged to do in its unconditional pledge. As detailed below, in my model the unconditional pledge will correspond to the Nash equilibrium (NE) without any interaction, i.e. what the country would do based on its own interest and taking the behavior of other countries as given, and the conditional pledge will be interpreted as a request of a share of the potential additional surplus created. Then, after all the pledges are received, and developed countries have pledged to contribute a given amount through art 9, additional negotiations will take place to determine the transfers finally offered to the developing country at the implementation phase (mainly through article 6). If the transfer resulting from this negotiation equals the one requested, fine. However, if the transfer coming out from the negotiation is lower, the government of our developing country needs to come back to its constituency and explain that they are performing the additional effort for less, i.e. that they obtained a smaller share. This has costs, which can be political or reputational (be it national or international reputation). Hence, in my model, obtaining a final share which is smaller than the one requested in the conditional pledge has cost.

In addition to this "pledge-and-implement" mechanism, the Paris Agreement contains long-term provisions. The rulebook and the Katowice package mentioned above contain provisions that will govern the pledge-and-implement mechanism in the long-term. In general, all the provisions included in the Paris Agreement under the so-called "ambition cycle", designed to increase the ambition of the agreement, are related to the long-term provision considered here. However, to be fair, the long-term provision considered below goes beyond anything agreed upon, but the goal is to analyze the role that these long-term provisions could play in the Paris Agreement if fully developed.

To include these long-term provisions in my model, in the second part of the paper I assume that countries can sign, or not, a long-term provision, that sets out some rules that govern future short-term interactions. I consider that this long-term provision is incomplete in a double sense[3]. First, countries can agree on abatement levels and associated transfers, but they cannot negotiate and sign agreements over investments (no provision in the Paris Agreement indicates how to reach the goals). To avoid confusion, I call agreements which are incomplete in this sense 'non-comprehensive'. Second, the long-term agreement on abate-

[3] Barrett (2002) uses the term 'incomplete' in a third sense, in the context of International Environmental Agreements, referring to agreements signed only by a subset of all the countries.

ment levels and associated transfers is complete if it is contingent on a stochastic variable that summarizes climate uncertainty, and is incomplete when it cannot depend upon the stochastic component. Real-world treaties are unlikely to include obligations that depend on the realization of a stochastic variable. Even over an issue such as climate change, where uncertainty is a key ingredient of the problem, countries are unlikely to agree on transfers and abatement efforts that depend on the severity of climate change. Moreover, the problem is so complex and multidimensional that it is probably not possible to write a meaningful agreement as a function of the state of nature. I call long-term provisions which are incomplete in this sense 'non-contingent'.

The analysis shows that the pledge-and-implement mechanism can implement the first-best in the short-term, and that the surplus is ultimately shared according to the relative importance of the cost-of-revoking the pledges (more precisely, countries with relatively larger costs-of-revoking obtain a larger share). However, without a long-term provision there is underinvestment. The reason for this is the standard hold-up problem, as countries will underinvest to improve their bargaining power during the pledge and implement phases. However, if it were possible to write a long-term agreement contingent on the state of nature the first-best would be attaint. This holds even if the long-term provision can be abandoned at any time if not profitable for all parts, as I will assume below (countries can always abandon treaties with no penalties, as Canada abandoned the Kyoto Protocol just before the end of the first commitment period). Thus, if the long-term provision could be made contingent on the state of nature it would be able to solve the hold-up problem, despite its weakness. However, as it is probably not possible to write a long-term provision that is truly contingent on the state of nature, I analyze also the role that a non-comprehensive and non-contingent agreement can have. Not surprisingly the results show that the first-best can no longer be attaint. Nevertheless, even this form of long-term provision brings the world closer to the first-best. Furthermore, the analysis shows that the problem lays not so much in the non-comprehensiveness of the long-term provision but more on the fact that it is non-contingent. Thus, the more dependent the terms of the long-term provision is on the state of nature finally realized (e.g. on the severity of the climate change observed), the closer the agreement will bring us to the first best.

A large body of literature on game-theoretic analyses of International Environmental Agreements (IEA) has emerged over the last three decades. A recent survey and a selection of the most relevant papers can be found in Finus and Caparrós (2015). The literature on IEA can be divided into several broad categories. First, we have literature based on the concept of the internal and external stability of a coalition (agreement), in the sense that no player wants to leave or join the coalition (Carraro and Siniscalco 1993; Barrett 1994). Results show that for identical players, only a very small number of players form a coalition, while asymmetries among players allow for larger coalitions. The second approach is based on cooperative game theory and, more precisely, on the core (Chander and Tulkens 1997). The objective is to set up a burden-sharing rule that is able to favor the cooperation of all, ensuring that the rule prevents any individual country, but also any sub-group of countries, from being interested in leaving the IEA. The two approaches just mentioned are static but there is also a large set of papers that have focused on the dynamic nature of the climate change problem (see the survey in Calvo and Rubio 2013). The papers in this third approach are characterized by using dynamic state-space games to analyze IEA, combined with one of the stability concepts discussed above. The fourth approach focuses on the negotiation process, a part from which the previous approaches abstract. Here the goal is to investigate the impact of different bargaining protocols on the final agreement (for a survey, see Caparrós 2016).

None of these papers is appropriate to model the process described above: a long-term provision followed by a series of interactions governed by the pledge-and-implement mechanism. In all these papers the focus was on a long-term burden sharing agreement, such as the Kyoto Protocol. In the long list of papers pioneered by Carraro and Siniscalco (1993) and Barrett (1994), countries determine their emissions jointly once they sign the agreement (all signatories essentially become one player). The papers following Chander and Tulkens (1997) assume that a supranational authority imposes abatement efforts and transfers, ensuring that all players, and all groups of players, are better off with the proposed burden sharing rule. The agreement is then enforced by a threat to go back to all-singleton behavior. Papers in the bargaining tradition (Caparrós et al., 2004) also assume that all agreements are binding or, alternatively, that defection of one country will bring all countries back to the all-singleton behavior.

There is also a growing literature applying experimental economics methods to analyze alternative IEA designs. Within this group, Barrett and Dannenberg (2016) present a laboratory experiment that analyzes the submission of INDCs by countries followed by a review, with no legal consequences, of the degree in which the pledges have been met (the so-called 'pledge-and-review' process).

In modeling terms, my analysis of the pledge-and-implement mechanism is closely related to the bargaining model with partial commitment introduced by Muthoo (1996), in the context of the Nash Bargaining Solution, and analyzed using Rubinstein's alternating-offers model in Leventoglu and Tarar (2005). Focusing on this part of my model, my contribution is to extend their model to more than two players. This is not a trivial task in a model with alternating offers as mine (using the NBS and extending Muthoo's approach would also not be trivial, as he proposes a graphical proof).

In terms of motivation, the contemporaneous paper by Harstad (2019) is the most related one, as both model the Paris Agreement. They also share that the bargaining phase is integrated into a multi-stage model that analyzes long-term investments. His model details more the dynamics, using a quadratic specification, although mine is arguable more stylized. However, the key difference is that Harstad assumes that pledges need to be accepted by other countries, and that countries need to submit a new pledge if at least one country rejects any of the pledges submitted (he does not differentiate between conditional and unconditional pledges). A repetition of the submission of pledges if they are not accepted by all the countries that contribute positively would be a potentially interesting modification of the mechanism proposed in the Paris Agreement, but it is not a feature of the current version. My model stays, I would argue, closer to the bargaining mechanism introduced by the Paris Agreement. However, as discussed below, to obtain the results highlighted in my theoretical analysis, the Paris Agreement would also need to be developed. In this sense, both papers are complementary.

Other precedents to my analysis of the long-term provisions can be found in Harstad (2016), and to a lesser extend in Harstad (2012), Battaglini and Harstad (2016) and Beccherle and Tirole (2011). Although in a different framework, Harstad (2012 and 2016) analyze incomplete contracts in the first sense discussed above, i.e. in the sense that countries can negotiate and sign agreements on emissions and not on investment (what I have called non-comprehensive). Battaglini and Harstad (2016) extend the analysis to include the role of participation and duration of the agreements. Harstad (2012) and Battaglini and Harstad (2016) analyze a deterministic framework, while Harstad (2016) includes a stochastic component, and is therefore more relevant for my analysis. However, the agreements that he analyzes are complete in the sense that they depend on the stochastic variable (contingent in the terminology introduced above). As discussed above, contingent agreements are difficult to implement for a problem as complex as climate change. Furthermore, in this paper his analysis is based on a quadratic specification and, although he considers long-term agreements, he does not consider the two-tier procedure described above (one long-term provision followed by a sequence of short-term interactions). Other relevant precedents can be found in Dutta and Radner (2004, 2009, 2012) and in Barrett (2002). The former focus on equilibrium selection in analyses that combine theory and simulations and the latter focuses on the role of alternative assumptions when analyzing the possibility to renegotiate agreements.

The part devoted to the long-term provision is also closely related to the vast literature on incomplete contracts (in particular Muthoo 1995, but also Hart and Moore 1988, Hart and Holstrom 1987 or Crawford 1988), and several of the results shown below are well known in that literature (see Segal and Whinston 2010 for a survey). This part of my model is basically a hold-up model, although it does not assume self-investments as most of these models do, using the terminology used, e.g., in Segal and Whinston (2010). On the contrary, the model has imperfect coalition investments, despite the fact that investments only reduce the cost function of the country investing. The reason is that, due to the public good nature of the abatement done in the last stage, when one country increases its investment it improves the disagreement point for the remaining countries. The model also differs from most of this literature by the fact that it considers a larger number of countries, and not only two as this literature typically does.

The rest of the paper is organized as follows. Section 2 presents the multilateral bargaining model with conditional and unconditional pledges. Section 3 integrates this bargaining model in a multi-stage model where countries invest in long-term assets before entering the bargaining phases. Section 4 discusses the results and concludes.

2 Pledge and implement bargaining

There are N countries which can provide abatement efforts to mitigate a public bad problem, such as climate change (i.e., abatement is a public good). There are two phases: a first phase where countries submit their pledges simultaneously (their NDC) and a second phase where the pledges are (eventually) implemented. In the first phase ('pledge' phase) countries submit simultaneously their conditional and their unconditional NDC pledges. Unconditional pledges are independent of the behavior of other countries. Conditional abatement efforts are additional abatement efforts which are conditional on the behavior of other countries or on transfers. If all countries submit only unconditional pledges, at the second phase each country decides independently its abatement efforts. If at least one country submits a conditional pledge, in the second phase ('implementation' phase) countries negotiate about the implementation of the conditional pledges. Pledges can be revoked but at some cost, as detailed below. That is, backing off from an initial pledge has costs and countries suffer a greater loss the greater the distance from their initial pledge to the agreement finally implemented. This cost can be seen as reputational (national or international reputation, see the introduction).

Gross payoffs are given by the function

$$v_i(\mathbf{q}) = B_i(Q) - C_i(q_i),\ \forall i \in N \tag{1}$$

where $\mathbf{q} = (q_1, ...q_N)$ shows the abatement effort finally done by each country, B_i summarizes benefits from climate change abatement and C_i shows the cost incurred through the abatement efforts. Benefits B_i depend on aggregate abatement ($Q = \sum_{i \in N} q_i$). Costs C_i depend on the amount of abatement performed by county i . The discount rate is given by $r > 0$, with $\delta \equiv e^{-r\tau}$. Regarding the payoff function mentioned above, I make the following standard assumptions:

Assumption 1.a: *For all* $i \in N$: $\frac{\partial B_i}{\partial q_i} > 0,\ \frac{\partial^2 B_i}{\partial q_i^2} \leq 0,\ \frac{\partial C_i}{\partial q_i} > 0,\ \frac{\partial^2 C_i}{\partial q_i^2} > 0, \lim_{Q \to 0} \frac{\partial B_i}{\partial q_i} > \lim_{q \to 0} \frac{\partial C_i}{\partial q_i} > 0.$

To provide a benchmark, let use denote the Nash equilibrium of the game without interactions, where each country solves $\arg\max_{q_i} v_i(q_i, \mathbf{q}_{-i})$, as $\mathbf{q}^A = (q_1^A, ...q_N^A)$. This is the equilibrium without any pledges and any negotiations. However, given assumption 1.a this implies a positive level of abatement for all countries.

As 'unconditional NDC pledges' are independent of the behavior of other countries, each country simply submits its own pledged abatement effort q_i^U. The vector of all pledges is noted: $\mathbf{q}^U = (q_1^U, ...q_N^U)$. Note that unconditional NDC are not accepted by other countries. A 'conditional NDC pledge' by country i is a pair formed by (1) an abatement vector, noted $\mathbf{q}^C = (q_1^C, ...q_N^C)$, where $q_i^C \geq q_i^A\ \forall\ i$ and at least for one country $q_j^C > q_j^A$, and (2) a requested share for country i of the additional surplus created equal to α_i. That is, for at least one country the abatement proposed in the conditional pledges is above the NE without interaction, and the country submitting the conditional pledge requests a share of the additional surplus created. For example, a conditional pledge from country i can request a transfer from country j in exchange of an additional abatement from country i, as proposed by several developing countries (see the introduction).

Before continuing, let us note first that the abatement effort included in the unconditional pledges is identical to one associated to the Nash equilibrium of the game without any negotiations, and that, on the aggregate, these abatement efforts fall short from the first-best abatement effort, noted $\mathbf{q}^* = (q_1^*, ...q_N^*)$. For future reference, the following Lemma highlights this straightforward result:

Lemma 1 *Abatement efforts included in the unconditional NDC are identical to those corresponding to the Nash Equilibrium of the game without interaction, i.e.* $\mathbf{q}^U = \mathbf{q}^A$. *These abatement efforts are smaller than the first-best efforts at the aggregate level, i.e.* $\sum_{i \in N} q_i^U < \sum_{i \in N} q_i^*$.

Proof. Each country solves $\arg\max_{q_i} \frac{v_i(\mathbf{q}, I_i, \theta)}{1-\delta}$ taking $\mathbf{q}_{-i}$ as given. This yields the standard Nash Equilibrium:

$$\frac{\partial B_i}{\partial q_i} - \frac{\partial C_i}{\partial q_i} = 0 \Longrightarrow q_i^U = q_i^A \quad \forall i \in N \tag{2}$$

First-best abatement efforts are given by

$$\sum_{i \in N} \frac{\partial B_i}{\partial q_i} - \frac{\partial C_i}{\partial q_i} = 0 \Longrightarrow q_i^* \quad \forall i \in N \tag{3}$$

and it is easy to show that they are larger at the aggregate level. ■

On the contrary, because conditional pledges imply, by definition, that at least one country will abate more if implemented, conditional pledges create an additional surplus. Let us denote this additional surplus by

$$\Delta(\mathbf{q}^A, \mathbf{q}^C) = \sum_{i \in N} \left[v_i(\mathbf{q}^C) - v_i(\mathbf{q}^A) \right].$$

As already noted, pledges can be revoked but at some cost. That is, backing off from an initial pledge has costs and countries suffer a greater loss the greater the distance from their initial pledge to the agreement finally implemented. To avoid unnecessary complications, I assume that unconditional pledges are never revoked (as Lemma (1) has shown, unconditional pledges just write down Nash equilibrium efforts). For conditional pledges, I assume the following functional form for the cost-of-revoking function:

$$K(\alpha_i, \varphi_i) = \begin{cases} 0 \text{ if } \alpha_i \leq \varphi_i \\ k_i(\alpha_i - \varphi_i) \text{ if } \alpha_i > \varphi_i \end{cases} \tag{4}$$

where $k_i > 0$ and φ_i is the share finally implemented. That is, countries suffer more the larger the distance between the share requested in their conditional NDC and the share finally obtained. To keep the model as simple as possible, I assume a linear relationship. If countries are unable to agree at the implementation phase, the disagreement point is given by the unconditional pledges.

At the implementation phase, countries either implement their unconditional abatement efforts (Nash equilibrium) or sign agreements based on articles 6 and 9 to implement the additional efforts included in the conditional pledges. These agreements are negotiated following a standard offer and counter-offer procedure, but taking into account that modifying the position set out in the pledges has cost. To simplify, I assume a cyclical protocol where the order in which countries talk is known (the order in which countries talk captures the relative importance of different countries in the international arena). I also assume that countries talk in the order given by their ordinality (as I can rename countries at will this implies no loss of generality).

Summing up, countries submit their pledges simultaneously and, if all countries submit only unconditional pledges, or if (some) countries submit conditional pledges but there is no agreement at the implementation phase, the net payoff for country i is given by:

$$\Pi_i(\alpha_i, \varphi_i, \mathbf{q}^A) = v_i(\mathbf{q}^A)$$

If (some) countries submit conditional pledges and there is an agreement at the implementation phase, the net payoff for country i is given by:

$$\Pi_i(\alpha_i, \varphi_i, \mathbf{q}^A, \mathbf{q}^C) = v_i(\mathbf{q}^A) + \varphi_i \Delta - K(\alpha_i, \varphi_i) \; \forall \; i \in N \tag{5}$$

Restricting attention to stationary strategies to simplify the analysis, the following proposition describes the equilibrium NDC pledges and the associated payoffs[4]:

Proposition 1 *Each country submits an unconditional NDC where q_i^U is defined by the Nash equilibrium of the game without interactions, as define in equation (1), i.e. $q_i^U = q_i^A$, and a conditional NDC consisting of a set of abatement efforts $\mathbf{q}^C = \mathbf{q}^*$ and associated transfers that imply a share for country i of the additional surplus created, $\Delta(\mathbf{q}^A, \mathbf{q}^C)$, equal to*

$$\alpha_i^* = \frac{\delta^{i-1}(1+k_i)}{\sum_{i=1}^{N} \delta^{i-1}(1+k_i)} \forall i \in N.$$

In the unique equilibrium, the first-best abatement efforts, $\mathbf{q}^C = \mathbf{q}^*$, are implemented and the shares finally implemented are $\varphi_i^* = \alpha_i^* \; \forall i \in N$, with $\sum_{i=1}^{N} \varphi_i^* = 1$. Thus, no country needs to back-up from its initial conditional NDC and net payoffs are given by:

$$\Pi_i = v_i(\mathbf{q}^A) + \varphi_i^* \Delta(\mathbf{q}^A, \mathbf{q}^*) \; \forall \; i \in N$$

as $K(\alpha_i^*, \varphi_i^*) = 0$ for $\alpha_i^* = \varphi_i^*$

Proof: Appendix A.1

As discussed above, countries agree on the abatement level that maximizes the additional surplus to be shared, and each country obtains a share φ_i of the additional surplus created, plus its payoff in the situation where all countries follow their unconditional NDC (the Nash equilibrium without interactions). Although the detailed proof can be found in the Appendix, the intuition that explains why countries implement the first-best abatement effort

[4] If countries do not use stationary strategies the multilateral bargaining game would yield multiple equilibria. There are different alternatives to overcome this difficulty, such as the "exit game" in Krishna and Serrano (1993) or the model in Huang (2002), which has been extended to games with externalities in Caparrós and Péreau (2017). Although the proof would be more complicated, dropping the assumption of stationary strategies and adapting Huang's model would probably yield similar results, as long as one assumes that the discount rate tends to one and each country has the right to talk once following a cyclical protocol (as the solution converges to the solution with stationary strategies under these assumptions).

is as follows. Any set of abatement efforts that falls short of the socially optimal abatement effort cannot be a Nash Equilibrium, because countries will be interested in increasing the ambition to avoid leaving unexhausted surplus. In equilibrium, conditional pledges can also not include a set of transfers that implies that the sum of the shares requested of the surplus created exceed one. The reason is that at least one country would be interested in requesting a smaller share, to avoid revoking costs. A set of transfers that implies that the sum of the shares requested is less than one can also not be an equilibrium, as at least one country would be interested in requesting a larger share. Thus, abatement efforts and requested shares lay in equilibrium on the Pareto frontier. What the Appendix shows is that only one point on the Pareto frontier can be sustained as a Nash equilibrium. In that point the additional surplus obtained by moving from the disagreement point (where all countries implement their unconditional NDC) to the situation where all implement their conditional NDC, is shared proportionally to the cost-of-revoking function of each country. Countries that suffer higher costs from modifying their initial pledges obtain a larger share of the surplus created. Nevertheless, these costs-of-revoking never actually occur in equilibrium, as in our world countries submit "perfect" conditional NDC pledges.

As already mentioned, the share obtained from the additional surplus depends on the relative cost for the country of modifying its position. If all have the same cost-of-revoking function, the split the difference rule holds and $\varphi_i = \varphi = 1/N \ \forall i$. If the cost for one country of changing its conditional contribution tends to infinity, while the costs of the remaining countries is finite, the uncompromising country obtains all the surplus. Note also that, the total abatement achieved with the "pledge and implement" procedure is first best, and identical to the one obtained using the asymmetric Nash bargaining solution. However, the distribution of the surplus generated is not identical.

3 Long term provisions and investment

The last section has shown that the "pledge and implement" mechanism can bring abatement levels to the first best in the short term. In this section I analyze the role of long-term provisions in the Paris Agreement. Long term provisions are relevant because through long-term investments countries can reduce the costs of short-term abatement efforts. In addition, I incorporate uncertainty in the model, a key feature of the climate change problem. As discussed in the introduction, the Paris and its recent developments contain several long-term provisions. However, as also discussed in the introduction, to be fair, the long-term provision considered in this section goes beyond anything agreed upon, but the goal is to analyze the role that these long-term provisions could play in the Paris Agreement if fully developed.

The extended model has four stages, as shown in Figure 1. At stage 1, countries decide to sign, or not, a long term 'non-comprehensive' provision. The long-term provision is non-comprehensive because it cannot specify investment levels. That is, the long-term provision only specifies the additional abatement that countries will perform at stage 4, and the associated transfers, as detailed below. Note that this is consistent with the fact that investments are not covered in any form in the Paris Agreement. In addition, this long-term provision can be 'contingent', if it is written as a function of the state of nature described below, or 'non-contingent', if it is not contingent on the state of nature.

[Figure 1 about here]

Whether or not a long-term provision is included in the Paris Agreement, countries decide at stage 2 their investment levels, denote by $\mathbf{I} = (I_1, ...I_N)$. Investing I_i units has a cost equal to I_i for country i. This investment is private and benefits directly only the country investing, by reducing its abatement cost in the subsequent stages.

At stage 3, the state of nature θ is randomly realized, out of the finite set of possible states of nature Θ. The probability of each state of nature is known. After the value of θ is observed at stage 3, it becomes common knowledge.

Stage 4 has an infinite number of identical periods. Each one of these periods can be divided into the two phases described in the previous section. The only relevant modification to the analysis shown above is that gross payoffs[5] per period are now given by the function

$$v_i(\mathbf{q}, I_i, \theta) = B_i(Q, \theta) - C_i(q_i, I_i),\ \forall i \in N.$$

This function is similar to the function introduced in equation (1) except that now benefits B_i depend on aggregate abatement ($Q = \sum_{i \in N} q_i$) and on the realization of a random variable θ, and costs C_i depend on the amount of abatement performed by county i at stage 4 and on the investment done by county i at stage 1 (investment stage). Investment is a private good, and only the country investing benefits from the costs reduction. In addition to Assumption 1.a, I make the following assumptions:

Assumption 1.b: *For all* $i \in N$: $\frac{\partial C_i}{\partial I_i} < 0$, $\frac{\partial^2 C_i}{\partial I_i^2} > 0$, $\frac{\partial^2 C_i}{\partial q_i \partial I} < 0$, $\frac{\partial B_i}{\partial \theta} > 0$.

Because all periods at stage 4 are identical and $\mathbf{I}$ and θ are already known at the beginning of stage 4, the analysis shown in section 2 is valid for any of the periods in stage 4 if no long term provision is signed at stage 1. That is, if no long-term provision is signed, countries submit their NDC pledges freely and the pledge-and-implement process is repeated in each period. If there is perpetual disagreement at the implementation phase in one period, countries follow the Nash equilibrium without interaction from then onwards (their unconditional pledges). If a long-term provision was signed, they follow its terms if they are profitable for all parts, taking into account the value of θ observed at stage 3, as detailed below. Otherwise, countries are free to send their NDC pledges freely again.

Before continuing, I define formally self-investments and cooperative investments for future reference. As already mentioned in the introduction, although investment is a private good in the sense that it benefits directly only the country that undertakes the investment, using the terminology commonly used in the literature on incomplete contracts (Segal and

[5] Note that none of the results shown below depend on the separable nature of benefits and costs, as I could work directly with the function v. Nevertheless, I use functions B_i and C_i to facilitate the interpretation and the comparability with the rest of the literature on IEA.

Whinston, 2010) it is a cooperative investment, in the sense that it affects the disagreement payoff of other countries.

Definition 1 *Country i's investment is a "self-investment" if country j's disagreement payoff,* $v_j(\mathbf{q}^A, I_j, \theta)$*, is independent of* I_i *for all* $j \neq i$*, at all* I_{-i} *and* θ*. Country i's investment is "cooperative" if country j's disagreement payoff is non decreasing in* I_i *for all* $j \neq i$*, at all* I_{-i} *and* θ*.*

3.1 First best

To provide a benchmark, let us analyze the first best outcome for this extended model.

Lemma 2 *The first best equilibrium in terms of abatement and investment is defined by the following equations:*

$$E_\theta \left[\sum_{i \in N} \frac{\partial B_i}{\partial q_i} - \frac{\partial C_i}{\partial q_i} \right] = 0 \Longrightarrow q_i^* = q_i^*(I_i, \theta) \quad \forall i \in N \tag{6}$$

$$E_\theta \left[- \left. \frac{\partial C_i}{\partial I_i} \right|_{q_i^*(I_i,\theta)} \right] = 1 - \delta \quad \Longrightarrow I_i^* \quad \forall i \in N \tag{7}$$

Proof. As all short term interactions are identical, the central planner has the following objective function:

$$f(\mathbf{q}, I_i, \theta) = \arg\max_{\mathbf{I},\mathbf{q}} \left(E_\theta \left[\frac{\sum_{i \in N} v_i(\mathbf{q}, I_i, \theta)}{1-\delta} \right] - \sum_{i \in \mathbf{I}} I_i \right) \tag{8}$$

The FOC for the abatement efforts are given by:

$$\frac{\partial f}{\partial q_i} = E_\theta \left[\frac{1}{(1-\delta)} \left(\sum_{i \in N} \frac{\partial B_i}{\partial q_i} - \frac{\partial C_i}{\partial q_i} \right) \right] = 0 \quad \Longrightarrow \quad q_i^* = q_i^*(I_i, \theta) \quad \forall i \in N, \tag{9}$$

yielding equation (6). Define the maximum value function:

$$F(\mathbf{q}, \mathbf{I}, \theta) = f(\mathbf{q}^*(\mathbf{I},\theta), \mathbf{I},\theta) = f(q_1^*(I_1,\theta), ..., q_N^*(I_N,\theta), I_1, ..., I_N, \theta)$$

where $q_i^*(I_i,\theta)$ are the optimal levels of abatement stage 4. From the envelope theorem, using (9):

$$\frac{\partial F(\mathbf{q},\mathbf{I},\theta)}{\partial I_1}=\frac{\partial f}{\partial q_1}\frac{\partial q_1}{\partial I_1}+...+\frac{\partial f}{\partial q_N}\frac{\partial q_N}{\partial I_1}+\frac{\partial f}{\partial I_1}=\frac{\partial f}{\partial I_1}$$

and the FOC in terms of investment is

$$E_\theta\left[\frac{\sum_{i\in N}\frac{\partial v_i(q_i^*(I_i,\theta),I_i,\theta)}{\partial I_i}}{1-\delta}\right]-1=E_\theta\left[\frac{-\left.\frac{\partial C_i}{\partial I_i}\right|_{q_i^*(I_i,\theta)}}{1-\delta}\right]-1=0 \quad \forall i\in N,$$

or equation (7). ■

The first order condition in (6) is identical to (3), just taking into account the role of I_i and θ. It is the standard equation defining optimal abatement, only taking into account the role of the investment made and that abatement benefits have a random component. Condition (7) yields the optimal investment level. The central planner takes into account the impact of the investment of country i on the infinite stream of future costs of abatement. However, as one only needs to take into account direct effects, due to the envelope theorem, and investments are a private good and only benefit directly the country that undertakes the investment, the FOC in terms of investment are independent from each other. Thus, each FOC yields the optimal investment for one country.

3.2 The investment stage in the "basic Paris Agreement"

Consider now the situation where there is no long-term provision determining the pledges that countries should submit in each period. To solve the game by backward induction, let us start by stage 4. In this case, in each period of stage 4 the situation is akin to the bargaining problem with two phases analyzed in section 2. However, the surplus created in each period depends now on the investment level and on θ:

$$\Delta(\mathbf{q}^A,\mathbf{q}^C,\mathbf{I},\theta)=\sum_{i\in N}\left[v_i(\mathbf{q}^C,I_i,\theta)-v_i(\mathbf{q}^A,I_i,\theta)\right].$$

That is, the share now depends on $\mathbf{I}$ and θ. Nevertheless, because both are known at the beginning of stage 4, the analysis shown in section 2 holds without any significant modification.

At stage 3 there are no strategic decisions (only θ is revealed at this stage). The following proposition summarizes the results concerning the investment level (stage 2), where I_i^S denotes investment without any long-term provision.

Proposition 2 *Countries will not implement the first best through a series of short term interactions (pledge-and-implement) without a long-term provision, as investment, I_i^S, will be defined, $\forall i \in N$ and $j \neq i$, by*

$$\underbrace{\underbrace{\varphi_i E_\theta \left[-\left.\frac{\partial C_i}{\partial I_i}\right|_{q_i^*(I_i,\theta)} \right]}_{\text{Conditional NDC}} + \underbrace{(1-\varphi_i) E_\theta \left[-\left.\frac{\partial C_i}{\partial I_i}\right|_{q_i^A(I_i,\theta)} \right]}_{\text{Unconditional NDC}}}_{\text{Self-investment terms}} \underbrace{- \varphi_i \sum_{j \in N/i} E_\theta \left[\left.\frac{\partial B_j}{\partial q_i}\right|_{q_i^A(I_i,\theta)} \frac{q_i^A(I_i,\theta)}{\partial I_i} \right]}_{\text{Cooperative investment term}} = 1-\delta \tag{10}$$

Thus, there will be underinvestment as compared to the first best and as a consequence abatement will also be below the optimal level. Overall underinvestment will be more severe the larger the number of countries, and the larger the benefit obtained by other countries from the increased abatement in case of disagreement induced by larger investments. For a particular country, a larger share of the surplus generated by the short-term agreement, φ_i, reduces underinvestment if increasing investment has a stronger impact on payoffs in case of agreement than in case of disagreement.

Proof. From Proposition 1 we know that at each interaction at stage 4 countries agree on the abatement level that maximizes the additional surplus to be shared. Taking into account the role of I_i and θ this implies that $q_i^C(I_i,\theta) = q_i^*(I_i,\theta)$ $\forall i \in N$ is given by (6). Thus, at stage 2 (investment stage), the expected payoff for each country i is:

$$E_\theta \left[\frac{v_i(\mathbf{q}^A, I_i, \theta) + \varphi_i \sum_{j \in N} \left[v_j(\mathbf{q}^*, I_j, \theta) - v_j(\mathbf{q}^A, I_j, \theta) \right]}{1-\delta} \right] - I_i \tag{11}$$

i.e., its disagreement payoff in each period plus a share of the additional surplus created through the pledge-and-implement bargaining process in each period, minus the cost associated with the investment. And the NE at stage 2 is given by:

$$E_\theta\left[\frac{\partial v_i(\mathbf{q}^A, I_i, \theta)}{\partial I_i}\right] + \varphi_i \sum_{j\in N} E_\theta\left[\frac{v_j(\mathbf{q}^*, I_j, \theta)}{\partial I_i} - \frac{\partial v_j(\mathbf{q}^A, I_j, \theta)}{\partial I_i}\right] = 1 - \delta$$

Applying again the envelope theorem, but taking into account that $\frac{\partial v_j(\mathbf{q}^A, I)}{\partial q_i^A} \neq 0$, we obtain (10). Overall underinvestment will be more severe the larger the number of countries because the last term on the LHS of (10) becomes larger, for a given φ_i. Larger φ_i reduces underinvestment if $E_\theta\left[-\left.\frac{\partial C_i}{\partial I_i}\right|_{q_i^*(I_i,\theta)}\right] > E_\theta\left[-\left.\frac{\partial C_i}{\partial I_i}\right|_{q_i^A(I_i,\theta)} + \sum_{j\in N/i} \left.\frac{\partial B_j}{\partial q_i}\right|_{q_i^A(I_i,\theta)} \frac{q_i^A(I_i,\theta)}{\partial I_i}\right]$, yielding the last statement in the proposition. ■

As the proposition shows, for a given country, increasing its investment at stage 2 has three effects. First, it decreases its costs in the event of an agreement at stage 4. This is the first term on the LHS of (10). As highlighted in equation (10), this term is related to the conditional NDC, as the derivative of the cost function is evaluated at the abatement level included in the conditional NDC. The larger the share of the surplus created by the agreement that the country appropriates, φ_i, the more important this term becomes. The second effect is that it reduces its costs in the event of a disagreement at stage 4. This is the second term on the LHS in (10) and the derivative of the cost function of this term is evaluated at the abatement level included in the unconditional NDC. The larger the share obtained by country i the less important this term becomes. The two terms just discussed appear if investment is a self-investment and if it is a cooperative investment (see Definition 1), as they refer to the impact that investment has on the agreement and disagreement payoffs of the country itself. The last term on the LHS appears only because investment is cooperative in this model. The investment done by country i at stage 2 impacts the disagreement point of all the other countries. The reason is that, although investment only impacts the cost function of the investing country, the investment implies more abatement by country i at stage 4 and this improves the disagreement point of all the other countries. This term becomes more important the larger the share of the surplus obtained by the country, and reduces the investment level of country i, as improving the disagreement point of the other countries reduces the bargaining power of the country investing.

Figure 2 illustrates equations (7) and (10). The thick horizontal line is the RHS of (7) and (10). The dashed line is the LHS of (7). The first-best investment level, I^*, is given by the intersection of these two lines. The thin solid line is the first term on the LHS of (10). Note that if investment had no impact on the disagreement point, the second and third terms of (10) would vanish, and I^S would be given by the intersection of both solid lines. In fact, this is the case in the standard model analyzed in Hart and Moore (1988) or Muthoo (1995), where the no-agreement situation implies no-trade between a buyer and seller. It were also the case if one would assume, to simplify the model, that there is no abatement in case of disagreement. The reduction would be explained exclusively by the share φ_i that the country receives from the additional surplus created, as the individual country only cares about the share of the additional surplus created through the investment that it appropriates. The line with full dots includes the first two terms on the LHS of (10). That is, the situation that would prevail if there is abatement at the disagreement point, and hence investment affects this disagreement point, but investment were self-investments and would not impact the disagreement point of other countries. As shown in the Figure, the impact on the own disagreement point brings the investment level closer to the first-best level. The reason is that there is now a second incentive to invest, to reduce the costs in the event of a disagreement. Finally, the line with empty dots includes all the terms on the LHS of (10). Including the 'cooperative investment' term moves the investment level again further away from to the first-best situation. As already mentioned, the reason is that increasing the investment level increases the benefits of all other countries at the disagreement point, and this weakens the bargaining position of the country investing. That is, although the hold-up problem still exists, it is less severe than in the canonical model where the no-agreement situation implies no-trade between a buyer and seller.

[**Figure 2 about here**]

The following example illustrates the model using particular functions that are similar to those usually used in the literature on IEA (Barrett, 1994).

Example 1 *Consider $N \geq 3$ identical countries, implying $\varphi_i = \varphi = 1/N$ $\forall i$. Payoffs are $B(Q,\theta) = b\theta Q$ and $C(q,I) = \frac{c}{2I}\left(q^2 + F\right)$, where b, c and F are non-negative parameters, with $E_\theta\left[N^2 b^2 \theta^2\right] < 2(1-\delta)c$. The states of nature in Φ are $\{0.5, 1, 1.5\}$, with equal probability. Linear benefits are used frequently in this literature, even in papers where they are assumed to capture the infinite stream of future benefits. In the context analyzed here, $B(Q,\theta)$ only captures benefits in one period, and a linear specification is therefore more appropriate. The abatement cost function is also similar to the one frequently used in this literature (identical for $I = 1$), although with an additional fix cost F. However, in this model investment at stage 1 can reduce abatement costs (for $I > 0$). Dropping subscripts, optimal abatement and investment levels at the first best are*

$$\begin{aligned} q^*(I,\theta) &= \frac{Nb\theta I^*}{c} \\ I^*(\theta) &= \left(\frac{Fc^2}{2c\left(1-\delta\right) - N^2 b^2 \theta^2}\right)^{\frac{1}{2}}. \end{aligned}$$

For the case of a series of pledge-and-implement interactions but a NE at the investment stage with no long-term provision, we have:

$$\begin{aligned} q^S(I,\theta) &= \frac{Nb\theta I^S}{c} \\ I^S(\theta) &= \left(\frac{Fc^2}{2c\left(1-\delta\right) + b^2\theta^2\left(N - 3 + 1/N\right)}\right)^{\frac{1}{2}}. \end{aligned}$$

As predicted, in this case we have underinvestment: $I^S < I^*$. To see that $I^S < I^*$, subtract the denominator from the expression for I^S from the denominator of the expression for I^*. This yields $\frac{1}{N} b^2 \theta^2 \left(N-1\right)\left(N^2 + 2N - 1\right)$, which is positive. Note also that $I^S > 0$ for $N \geq 3$ and that we also have $q^S < q^*$. Figure 2 is drawn using the functional forms in this example, with $(N, b, \theta, c, F, \delta) = (3, 1, 1, 50, 3, 0.9)$. For these numbers, in terms of investment we have $I^* = 87$ and $I^S = 31$, and in terms of abatement $q^* = 5.1962$ and

$q^S = 1.8766$. *That is, although the expressions for abatement are similar in both cases, underinvestment implies a significantly lower level of abatement if there is no long-term provision. If there is no long-term provision, abatement at the disagreement point would be* $q^A(I^S, 1) = 0.6255$ *while abatement at the disagreement point with the first-best investment level would be* $q^A(I^*, 1) = 1.7321$.

3.3 Non-comprehensive but contingent long-term provision

I now assume that countries include a long-term provision in the Paris Agreement, written as a function of θ. This long-term provision, written at stage 1, before θ is known, specifies the transfers that each country will receive and the abatement that it will perform in each short-term interaction implemented at stage 4. This long-term provision is not binding in the sense that countries can break it. Note that I am here referring to provisions included in the Paris Agreement, which is in general not binding, not at the potential shot-term contracts written by developing articles 6 and 9 of the Paris Agreement. If countries want to renegotiate, they simply abandon the long-term provision and negotiate freely following the pledges and implementation phases described in section 2. To simplify the analysis I assume that writing a new long-term provision after stage 1 is not possible. This can be justified by the effort needed to negotiate a new long-term agreement as the Paris Agreement. Note, however, that this is not really a relevant restriction as a new long-term provision after investment was made at stage 2 and uncertainty was revealed at stage 3 would not add any strategic component to the sequence of short-term interactions.

Let us fix an arbitrary long-term provision $\langle \tau(\theta) \rangle_{\theta \in \Theta}$ such that for any θ in each short interaction at stage 4:

$$v_i(\mathbf{q}^L, I_i, \theta) + \tau_i(\theta) \geq v_i(\mathbf{q}^A, I_i, \theta) \quad \forall i \in N, \tag{12}$$

where L in superscript stands for contingent long-term agreement. By analogy to the analysis in section 2, we know that abatement in case of agreement is optimal for a given level of investment, as countries have no interest in leaving unexhausted surplus, i.e. $q^L(I_i,\theta) = q^*(I_i,\theta)$. Given my assumptions on v_i there are many long-term provisions that satisfy the restrictions in (12). Let us further assume that the terms of the long-term provision will only be abandoned if it is not profitable for all parts, in the sense that at least one[6] country can credibly threat to prefer the situation that would prevail without any short-term interactions. Abandoning the long-term provision implies no punishment, but it implies to re-negotiate abatement efforts. As the outcome of this re-negotiation may, in principle, imply no agreement, at least one country should be able to threat that it prefers the situation without any agreement. That is, if (12) were to fail, the pledge and implement mechanism would go back to the situation without a long-term provision. However, unlike in the situation analyzed in the next section, there are many long-term provisions for which condition (12) would never fail, and I assume that $\langle\tau(\theta)\rangle_{\theta\in\Theta}$ is such a provision.

That is, the contingent long-term provision $\langle\tau(\theta)\rangle_{\theta\in\Theta}$ defined above will never be denounced, as it ensures that (12) holds for any realization of θ. Thus, the expected payoff for country i for any $\mathbf{I}$ at stage 2 is:

$$E_\theta\left[\frac{v_i(\mathbf{q}^*, I_i, \theta) + \tau_i(\theta)}{1-\delta}\right] - I_i \quad \forall i \in N \tag{13}$$

Hence, the NE at the investment stage is given by

$$E_\theta\left[\frac{\partial v_i(\mathbf{q}^*, I_i, \theta)}{\partial I_i}\right] = 1-\delta \quad \forall i \in N$$

or

$$E_\theta\left[-\left.\frac{\partial C_i}{\partial I_i}\right|_{q_i^*(I_i,\theta)}\right] = 1-\delta \;\; \forall i \in N \tag{14}$$

which are the same as first-best FOCs (see equation (7)).

Thus, we can write:

Proposition 3 *Countries will implement the first best, in terms of abatement and in terms*

[6] Note that the long-term provision would be part of the Paris Agreement, or its developments (such as the Katowice package, see the introduction) and, unlike the pledges, would therefore be subject to the unanimity rule.

of investment, through a series of short-term interactions regulated by a non-comprehensive but contingent long-term provision. That is, even if countries cannot sign an agreement on investments the first best will be attaint if the long-term provision defines abatement efforts and transfers that are contingent on the state of nature.

The second and the third terms in (10) disappear because now the share obtained by each country of the additional surplus created through an agreement at stage 4 is not anymore under discussion at that stage, because transfers were already determined at stage 1, when the long-term agreement was signed. This implies that, despite the weakness of the long-term agreement, which can be abandoned at any time, the first best is attaint even if countries cannot agree on investments.

As discussed in the introduction, the Paris Agreement is incomplete in a double sense. First, it is non-comprehensive, as it does not determine investment levels. What Proposition 3 tells us is that this incompleteness is not problematic, as the world could attain the first-best by writing a long-term provision that is contingent on the state of nature. In theory, a detailed development of the rulebook, and articles 6 and 9 of the Paris Agreement, could make the long-term provision contingent on the state of nature. However, it is no secret that the practical challenges of this task would be enormous. This is the reason why I now move on to analyze a long-term provision that is also incomplete in this second sense, i.e. non-contingent.

3.4 Non-comprehensive and non-contingent long-term provision

Assume now that θ is so complex and multidimensional that it is not possible to write a meaningful long-term provision as a function of the state of nature. This is clearly the case for climate change, as the number of variables is extremely large and the uncertainties surrounding them are also large. Thus, countries cannot write down future abatement efforts and associated transfers as a function of the state of nature finally realized.

Let us therefore assume that countries write a long-term provision where abatement efforts and associated transfers are specified independently from the state of nature. These provision states, e.g., that country j has to pay country i a given amount of money for increasing its abatement efforts to a certain level. The difference with the previous section is that the amount paid is determined a priori, and that it does not depend on the state of nature finally realized.

Under this assumption, let us fix an arbitrary incomplete long-term provision $\tau = (\tau_1, ..., \tau_N)$. At stage 4, investment levels $\mathbf{I}$ are sunk and the state of nature θ is known. Once θ is observed by the parties, all countries are happy with the terms of this agreement if

$$v_i(\mathbf{q}^l, I_i, \theta) + \tau_i \geq v_i(\mathbf{q}^A, I_i, \theta) \quad \forall i \in N, \tag{15}$$

where l in superscript stands for long-term non-comprehensive and non-contingent agreement. If these conditions hold for all countries, I assume that they honour the agreement, as in the previous sub-section. If (15) does not hold for at least one country (an eventuality that was not possible in the previous sub-section), the long-term agreement will not be honoured and countries will renegotiate each pledge-and-implement short-term interaction following the two-phases process described in section 2. As before, I assume that countries cannot write a new long-term provision (this assumption has again no strategic bite as the state of nature is known and investments are sunk). That is, in each pledge-and-implement short-term interaction countries will renegotiate and agree on a share that implies a set of transfers τ' such that

$$v_i(\mathbf{q}^l, I_i, \theta) + \boldsymbol{\tau}'_i \geq v_i(\mathbf{q}^A, I_i, \theta) \text{ for } i = 1, ..N.$$

Note that this new sequence of short-term interactions has an expected payoff identical to (11). Hence, at stage 4 equilibrium expected payoffs are defined by the terms given by the long-term provision if they are profitable for all parts, and by the expected payoffs associated with the sequence of short-term interactions without any long-term provision otherwise. That is, the payoff is given by:

$$\begin{cases} \frac{v_i(\mathbf{q}^l, I_i, \theta)+\tau_i}{1-\delta} & \text{if } \tau_i \geq v_i(\mathbf{q}^A, I_i, \theta) - v_i(\mathbf{q}^l, I_i, \theta) \text{ for } i = 1, ..N \\ \frac{v_i(\mathbf{q}^A, I_i, \theta)+\varphi_i \sum_{i\in N}\left[v_i(\mathbf{q}^S, I_i, \theta)-v_i(\mathbf{q}^A, I_i, \theta)\right]}{1-\delta} & \text{otherwise.} \end{cases}$$

Define now, for each vector of investment $\mathbf{I}$, the set

$$\Psi(\mathbf{I}) = \left\{\theta \in \Theta : \ \tau_i \geq v_i(\mathbf{q}^A, I_i, \theta) - v_i(\mathbf{q}^l, I_i, \theta)\right\}.$$

Set $\Psi(\mathbf{I})$ includes all the states of nature, given the investment vector decided at stage 2, for which (15) holds. Hence, given the arbitrary long-term provision $\boldsymbol{\tau}$, the expected payoff at stage 2 for country i is:

$$E_{\theta\in\Psi(\mathbf{I})}\left[\frac{v_i(\mathbf{q}^l, I_i, \theta) + \tau_i}{1-\delta}\right] + E_{\theta\in\Theta/\Psi(\mathbf{I})}\left[\frac{v_i(\mathbf{q}^A, I_i, \theta) + \varphi_i \sum_{i\in N}\left[v_i(\mathbf{q}^S, I_i, \theta) - v_i(\mathbf{q}^A, I_i, \theta)\right]}{1-\delta}\right] - I_i \tag{16}$$

For abatement the FOC will again be given by (6), implying $q_i^l(I_i, \theta) = q_i^S(I_i, \theta) = q_i^*(I_i, \theta)$. For investment, applying again the envelope theorem, but taking into account that $\frac{\partial v_j(\mathbf{q}^A, I)}{\partial q_i^A} \neq 0$, the NE investment vector $\mathbf{I}^l$ is given by:

$$\begin{aligned} & E_{\theta\in\Psi(I)}\left[-\left.\frac{\partial C_i}{\partial I_i}\right|_{q_i^*(I_i,\theta)}\right] \qquad (17) \\ & +E_{\theta\in\Theta/\Psi(I)}\left[-\varphi_i \left.\frac{\partial C_i}{\partial I_i}\right|_{q_i^*(I_i,\theta)} - (1-\varphi_i)\left.\frac{\partial C_i}{\partial I_i}\right|_{q_i^A(I_i,\theta)} - \sum_{j\in N/i}\varphi_i \left.\frac{\partial B_j}{\partial q_i}\right|_{q_i^A(I_i,\theta)} \frac{q_i^A(I_i)}{\partial I_i}\right] \\ = \ & 1-\delta \quad \text{for } i = 1, ..N \text{ and } j \neq i \end{aligned}$$

Note that the LHS of (17) is a linear combination of (7) and (10), implying that the non-comprehensive and incomplete long-term provision brings the investment level closer to the first best. Thus, we can write the following proposition:

Proposition 4 *Countries will not implement the first best through a series of short term interactions regulated by a non-comprehensive and non-contingent long-term provision. Given a level of investment, abatement will be optimal, as it will be defined by (6), i.e.* $q_i^l(I_i, \theta) =$

$q_i^S(I_i, \theta) = q_i^*(I_i, \theta)$. *However, there will be underinvestment, although it will be less severe than in the absence of the long-term provision. As a consequence, abatement will also be below the optimal level but closer to the first best than without the long-term provision. Furthermore, each additional state of nature included in the long-term provision, i.e. each additional state of nature included in* $\Psi(I)$*, weakly approaches the solution to the first best. If* $\Psi(I) \to \Theta$ *a series of short term agreements following the incomplete long-term provision tends to the first best.*

As discussed above, long-term provisions in the Paris Agreement are, at best, non-comprehensive and non-contingent. However, despite this weakness, Proposition 4 shows that a long-term provision has a role to play, although it will not be able to bring investment to the optimal level, it will bring investments closer to the first best. Propositions 3 and 4 also tell us that being non-comprehensive is not the main weakness of the long-term provision, as investments would be optimal even if countries cannot agree on them, if they were able to write a contingent long-term provision, in the sense that all possible states of nature are considered. Furthermore, the more complete the long-term provision becomes in this sense, the closer it moves us to the first best. Thus, the Paris Agreement should define as clearly as possible the conditions under which transfers will occur under Articles 6 and 9 and these conditions should depend as much as possible on the severity of the climate change finally observed (and on other stochastic variables).

4 Discussion and conclusion

This paper presents a stylized model to analyze the two two-tier process initiated by the Paris Agreement, where a long-term non-binding agreement is followed by a series of short-term interactions governed by a pledge-and-implement mechanism. The pledge-and-implement mechanism has been modeled as a bargaining game in which countries submit conditional and unconditional pledges, which are not binding and not subject to the approval by other countries. Then, eventually, countries bargain over the conditions set out in the conditional pledges. However, backing-off from the initial terms proposed in the conditional pledges has costs.

The paper also analyzes long-term climate provisions that are incomplete in a twofold sense. In the first one, they are non-comprehensive because countries do not agree on investments, only on rules governing future short-term interactions on abatement levels and associated transfers. In the second sense, long-term provisions on abatement are incomplete because the climate change problem is so complex and multidimensional that it is not possible to write a meaningful agreement as a function of the state of nature. Hence, transfers are non-contingent and cannot depend on the state of nature finally observed.

Results show that the pledge and implement mechanism can implement the first-best in the short-term, and that the surplus is shared according to the relative importance of the cost-of-revoking the pledges. However, without a long-term provision there is underinvestment. Even if the agreement does not cover investments, a long term provision contingent on the state of nature can bring the world to the long-term first-best, and a non-contingent provision can bring it closer.

The result that the pledge-and-implement mechanism leads essentially to the same results as applying directly the Nash Bargaining Solution, could be seen as implying that the Paris Agreement is essentially a burden sharing mechanism. However, this is only true if all countries use conditional NDC and articles 6 and 9 are fully developed, allowing for a meaningful implementation phase. This is not the case in reality, so that the results obtained should be seen as an indication about where the Paris Agreement process should be directed towards. This result also shows that it was probably a wise strategy to move from negotiating directly a burden-sharing agreement (the approached followed for the Kyoto) to this more politically palatable process where countries submit freely their pledges and negotiations are only needed to modify, at some cost, the initial positions set out in the pledges.

My interpretation of the current situation of the Paris Agreement is as follows. Developing countries have submitted conditional pledges, in addition to their unconditional pledges. As shown above, this is an appropriate strategy as it can potentially bring us closer to the short-term first-best, for a given level of investment. Developed countries have submitted, strictly speaking, only unconditional pledges. However, as they have also, or will soon (in principle by 2025, see the introduction), submit pledges for their contribution to the transfers to developing countries, one could see them also as (partial) conditional pledges. In any case, developed countries should specify clearer the transfers that they are willing to make, and the additional efforts that they would require in exchange. Furthermore, in the Cancun Agreement the EU submitted another form of conditional pledge, stating a more ambitious target if "other developed" countries contribute in a similar manner. My analysis shows that it was probably a bad idea to abandon this path in the Paris Agreement and that future submission should focus more on conditional pledges.

In any case, even if the pledge-and-implement were fully developed, it cannot prevent the long-term underinvestment problem. The long-term provisions in the current version of the Paris Agreement are not sufficiently developed to qualify as a long-term provision in the sense discussed above. Although there is a long-term goal, and there is a floor for the intended transfers from developed countries (USD 100 billion a year), countries have not determined sufficiently future pledges and associated transfers, neither without (non-contingent) nor with (contingent) a reference to the underlying uncertainty. This is problematic, as it implies that long-term investments will not be optimal according to my analysis. Recent developments, such as the Paris rule-book and the Katowice package are steps in the right direction, but the condition for future transfers should be detailed much more to deter incentives to hold-up investments. The analysis has also shown that being non-comprehensive is not the main weakness of a potential long-term provision, as investments would be optimal even if countries cannot agree on them, as long as were able to write a contingent agreement, in the sense that all possible states of nature are considered. Furthermore, even "real world" long-term provisions, which by definition will not be fully contingent, can bring the world closer to the first-best and should therefore be pursued.

Taking my short and long-term results together, the bottom line is that, if fully developed, the Paris Agreement could turn out to be a good instrument. However, the analysis also shows that the current Paris Agreement needs to be further developed.

References

[1] Barrett, S. 1994. Self-enforcing international environmental agreements. *Oxford Econ Pap* 46: 878-894.

[2] Barrett, S., 2002. Consensus Treaties. *J Inst Theor Econ* 158: 529-547.

[3] Barrett, S. and Dannenberg, A., 2016. An experimental investigation into 'pledge and review' in climate negotiations. *Clim change* 138: 339-351.

[4] Battaglini, M. and Harstad, B., 2016. Participation and Duration of Environmental Agreements. *J Political Econ* 124(1): 160-204.

[5] Beccherle, J. and Tirole, J., 2011. Regional Initiatives and the Cost of Delaying Binding Climate Change Agreements. *J Public Econ* 95 (11-12): 1339-1348.

[6] Calvo, E. and Rubio, S.J., 2013. Dynamic Models of International Environmental Agreements: A Differential Game Approach. *Int Rev Environ Resour Econ* 6: 289–339.

[7] Caparrós, A., 2016. Bargaining and International Environmental Agreements. *Environ Resour Econ* 65(1): 5–31.

[8] Caparrós, A. and Péreau, J.C., 2017. Multilateral versus sequential negotiations over climate change. *Oxford Econ Pap* 69(2): 365–387.

[9] Caparrós, A., Péreau, J.-C. and Tazdaït, T., 2004. North-South Climate Change Negotiations: a Sequential Game with Asymmetric Information. *Public Choice* 121 (3-4): 455-480.

[10] Carraro, C. and Siniscalco, D., 1993. Strategies for the international protection of the environment. *J Public Econ* 52(3): 309-328.

[11] Chander, P. and Tulkens, H., 1997. The Core of an Economy with Multilateral Environmental Externalities. *Int J Game Theory* 26: 379-401.

[12] Crawford, V., 1988. Long-Term Relationships Governed by Short-Term Contracts. *Am Econ Rev* 78: 485-499

[13] Dutta, P.K. and Radner, R. 2004. Self-enforcing Climate-Change Treaties. *Proceedings of the National Academy of Sciences* 101: 4746–4751.

[14] Dutta, P.K. and Radner, R., 2009. A Strategic Analysis of Global Warming: Theory and Some Numbers. *J Econ Behavr Organ* 71: 187–209.

[15] Dutta, P.K. and Radner, R. 2012. Capital Growth in a Global Warming Model: Will China and India Sign a Climate Treaty? *Econ Theory* 49: 411–443.

[16] Finus, M. and Caparrós, A., 2015. *Game Theory and International Environmental Cooperation: Essential Readings*. Edward Elgar, Cheltenham.

[17] Harstad, B., 2012. Climate Contracts: A Game of Emissions, Investments, Negotiations, and Renegotiations. *Rev Econ Stud* 79(4): 1527-57.

[18] Harstad, B., 2016. The Dynamics of Climate Agreements. *J Euro Econ Assoc* 14(3): 719–752.

[19] Harstad, B., 2019. Pledge-and-Review Bargaining. Mimeo.

[20] Hart, O. and Holstrom, B., 1987. The Theory of Contract. In T. Bewley (ed.) *Advances in Economic Theory*. Cambridge University Press.

[21] Hart, O. and Moore., 1988. Incomplete Contracts and Renegotiation. *Econometrica* 56: 755-786.

[22] Huang, C.-Y., 2002. Multilateral bargaining: conditional and unconditional offers. *Econ Theory* 20: 401–412.

[23] Krishna, V. and Serrano, R., 1996. Multilateral bargaining. *Rev Econ Stud* 63:61–80.

[24] Leventoglu, B. and Tarar, A., 2005. Prenegotiation Public Commitment in Domestic and International Bargaining. Am Polit Sci Rev 99(3): 419-433.

[25] Muthoo, A., 1995. Bargaining in a Long-Term Relationship with Endogenous Termination. *J Econ Theory* 66: 590-598.

[26] Muthoo, A., 1996. A Bargaining Model Based on the Commitment Tactic. *J Econ Theory 69: 134-152.*

[27] Rogelj, J., Elzen, M., Höhne, N., Fransen, T., Fekete, H., Winkler, H., Schaeffer, R., Sha, F., Riahi, K. and Meinshausen, M., 2016. Paris Agreement climate proposals need a boost to keep warming well below 2 °C. *Nature* 534: 631-639.

[28] Segal, I. and Whinston, M. D., 2010. Property Rights. In Gibbons, R. and Roberts, J. (eds) *Handbook of Organizational Economics*. Princeton University Press.

[29] Stavins, R.N. and Stowe, R.C. (editors), 2017. *Market Mechanisms and the Paris Agreement*. Working Paper of the Harvard Project on Climate Agreements. Cambridge, Mass.

[30] UNFCCC, 2015. *Adoption of the Paris Agreement* (FCCC/CP/2015/L.9/Rev. 1). Available at unfccc.int.

A Proof of Proposition 1

From Lemma 1 we know that unconditional pledges are defined by the abatement level of the Nash equilibrium without interactions, i.e. $\mathbf{q}^U = \mathbf{q}^A$. We also know that equilibrium abatement efforts in the eventuality of an agreement based on the conditional pledges are first-best, i.e. $\mathbf{q}^C = \mathbf{q}^*$, as countries have no interest in leaving unexhausted surplus in the case of an agreement, i.e. the additional surplus created is $\Delta = \Delta(\mathbf{q}^A, \mathbf{q}^*)$. I focus in the remainder of the proof in the case with three players. Nevertheless, the same strategy can be applied to any number of countries. Thus, I label countries as 1, 2 and 3.

In any equilibrium we must have that $\sum_{j=1}^{3} \varphi_i = 1$. If the shares add up to more than one the equilibrium cannot be implemented and if they add to less than one part of the surplus would not be appropriated by any country. Nevertheless, note that there is no obvious reason why $\sum_{j=1}^{3} \alpha_i$ should add up to one, although I will show below that this is the case in equilibrium. At the implementation phase the conditional NDCs are already know, so that α_1, α_2 and α_3 are given. Equilibrium offers at the implementation phase have to meet the following relationships. At time 3, country 1 offers a share φ_1^3 for itself, a share φ_2^3 for country 3 and a share $(1 - \varphi_1^3 - \varphi_2^3)$ for country 3 (superscripts indicate the period when the offer is made). This provides a payoff equal to $\varphi_1^3 - k_1(\alpha_1 - \varphi_1^3)$ for country 1, equal to $\varphi_2^3 - k_2(\alpha_2 - \varphi_2^3)$ for country 2 and equal to $(1 - \varphi_1^3 - \varphi_2^3) - k_3(\alpha_3 - (1 - \varphi_1^3 - \varphi_2^3))$ for country 3. At the previous period (time 2), country 3 makes an offer that makes countries 1 and 2 indifferent between accepting it or waiting one period for the offer of country 1. Thus, the share offered to country 1, denoted φ_1^2, has to ensure a payoff for 1 equal to $\delta\left(\varphi_1^3 - k_1(\alpha_1 - \varphi_1^3)\right)$. The share offered to country 2 by country 3, φ_2^2, has to provide a payoff equal to $\delta\left(\varphi_2^3 - k_2(\alpha_2 - \varphi_2^3)\right)$. This leaves country 3 with a share $1 - \varphi_1^2 - \varphi_2^2$ and a payoff $(1 - \varphi_1^2 - \varphi_2^2) - k_3(\alpha_3 - (1 - \varphi_1^2 - \varphi_2^2))$. At time 1, country 2 makes an offer that leaves 1 and 3 indifferent between accepting or waiting for the next period equilibrium offer. That is, country 2 offers them a share φ_1^1 and φ_3^1, respectively, yielding payoffs equal to $\delta^2\left(\varphi_1^3 - k_1(\alpha_1 - \varphi_1^3)\right)$ and $\delta\left((1 - \varphi_1^2 - \varphi_2^2) - k_3(\alpha_3 - (1 - \varphi_1^2 - \varphi_2^2))\right)$ for countries 1 and 3. This leaves a share equal to $1-\varphi_1^1-\varphi_3^1$ and a payoff equal to $(1-\varphi_1^1-\varphi_3^1)-k_2(\alpha_2-(1-\varphi_1^1-\varphi_3^1))$ for country 2. Finally, at time 0 country 1 makes an offer that leaves 2 and 3 indifferent between accepting or rejecting. That is, country 1 offers shares φ_2^0 and φ_3^0, yielding a payoff of $\delta\left((1-\varphi_1^1-\varphi_3^1)-k_2(\alpha_2-(1-\varphi_1^1-\varphi_3^1))\right)$ and $\delta^2\left((1-\varphi_1^2-\varphi_2^2)-k_3(\alpha_3-(1-\varphi_1^2-\varphi_2^2))\right)$ for countries 2 and 3, respectively. This leaves country 1 with a share equal to $1-\varphi_2^0-\varphi_3^0$ and a payoff equal to $(1-\varphi_2^0-\varphi_3^0)-k_1(\alpha_1-(1-\varphi_2^0-\varphi_3^0))$.

Assuming stationarity, we know that the offers made by country 1 at time 0 and time 3 have to be identical, i.e. $\varphi_1^0=\varphi_1^3=\varphi_1^*$ and yield identical payoffs for countries 2 and 3, i.e.:

$$\begin{aligned}
\varphi_2^3-k_2(\alpha_2-\varphi_2^3) &= \delta\left(\left(1-\varphi_1^1-\varphi_3^1\right)(1+k_2)-k_2\alpha_2\right) \\
\left(1-\varphi_1^3-\varphi_2^3\right)(1+k_3)-k_3\alpha_3 &= \delta^2\left(\left(1-\varphi_1^2-\varphi_2^2\right)(1+k_3)-k_3\alpha_3\right)
\end{aligned} \tag{18}$$

From the discussion above we also know that the offers φ_1^2, φ_2^2, φ_1^1 and φ_3^1 have to ensure that the following equalities hold:

$$\begin{aligned}
\delta(\varphi_1^3-k_1(\alpha_1-\varphi_1^3)) &= \varphi_1^2-k_1(\alpha_1-\varphi_1^2) \\
\delta\left(\varphi_2^3-k_2(\alpha_2-\varphi_2^3)\right) &= \varphi_2^2-k_2(\alpha_2-\varphi_2^2) \\
\delta^2\left(\varphi_1^3-k_1(\alpha_1-\varphi_1^3)\right) &= \varphi_1^1-k_1(\alpha_1-\varphi_1^1) \\
\delta\left(\left(1-\varphi_1^2-\varphi_2^2\right)-k_3(\alpha_3-\left(1-\varphi_1^2-\varphi_2^2\right))\right) &= \varphi_3^1-k_3(\alpha_3-\varphi_3^1)
\end{aligned} \tag{19}$$

Substituting (19) into (18) yields

$$\varphi_1^*=\frac{\begin{array}{c}1+\left(1+\alpha_1\left(\delta+\delta^2\right)\right)k_1+(1-\alpha_2)k_2+(1-\alpha_3)k_3 \\ +\left(1-\alpha_2+\alpha_1\left(\delta+\delta^2\right)\right)k_1k_2+\left(1+\alpha_1\left(\delta+\delta^2\right)-\alpha_3\right)k_1k_3 \\ +(1-\alpha_2-\alpha_3)k_2k_3+\left(1-\alpha_2-\alpha_3+\alpha_1\left(\delta+\delta^2\right)\right)k_1k_2k_3\end{array}}{(k_1+1)(k_2+1)(k_3+1)\left(\delta^2+\delta+1\right)} \tag{20}$$

$$\varphi_2^* = \frac{\begin{array}{c}\delta + (\delta - \alpha_1\delta)\,k_1 + \left(\alpha_2 + \delta + \alpha_2\delta^2\right)k_2 + (\delta - \alpha_3\delta)\,k_3 \\ + \left(\delta - \alpha_1\delta + \alpha_2\left(1+\delta^2\right)\right)k_1k_2 + (\delta - \alpha_1\delta - \alpha_3\delta)\,k_1k_3 \\ + \left(\delta - \alpha_3\delta + \alpha_2\left(1+\delta^2\right)\right)k_2k_3 + \left(\delta - \alpha_1\delta - \alpha_3\delta + \alpha_2\left(1+\delta^2\right)\right)k_1k_2k_3\end{array}}{(k_1+1)\,(k_2+1)\,(k_3+1)\left(\delta^2+\delta+1\right)} \tag{21}$$

To be an equilibrium, for these values it must be true that $\varphi_1^* \leq \alpha_1$, $\varphi_2^* \leq \alpha_2$ and $1 - \varphi_1^* - \varphi_2^* \leq \alpha_3$. First, I show that there is an equilibrium where these inequalities become equalities. Then I show that this is the unique equilibrium. The system formed by (20)-(21) and $\varphi_1 = \alpha_1$, $\varphi_2 = \alpha_2$ and $1 - \varphi_1 - \varphi_2 = \alpha_3$ has the following solution:

$$\begin{aligned}\alpha_1 &= \varphi_1 = \frac{(1+k_1)}{(1+k_1) + \delta\,(1+k_2) + \delta^2\,(1+k_3)} = \frac{(1+k_1)}{\sum_{i=1}^{3}\delta^{i-1}\,(1+k_i)} \\ \alpha_2 &= \varphi_2 = \frac{\delta\,(1+k_2)}{(1+k_1) + \delta\,(1+k_2) + \delta^2\,(1+k_3)} = \frac{\delta\,(1+k_2)}{\sum_{i=1}^{3}\delta^{i-1}\,(1+k_i)} \\ \alpha_3 &= 1 - \varphi_1 - \varphi_2 = \frac{\delta^2\,(1+k_3)}{(1+k_1) + \delta\,(1+k_2) + \delta^2\,(1+k_3)} = \frac{\delta^2\,(1+k_3)}{\sum_{i=1}^{3}\delta^{i-1}\,(1+k_i)}\end{aligned} \tag{22}$$

To see that a larger α_1 cannot be chosen by country 1 substitute (20)-(21) into the payoff function (5) to obtain:

$$\Pi_1(\alpha_1, \varphi_1) = \Omega - \frac{k_1\Delta}{\delta^2+\delta+1}\alpha_1,$$

where Ω collects the terms that depend on $\delta, \alpha_2, \alpha_3, d, v_1(q^A), k_1, k_2$ and k_3 but are independent from α_1.

As can be seen, the payoff is a decreasing function of α_1 and country 1 is therefore not interested in proposing a lager α_1. I now show, a contrario, that country 1 is also not interested in proposing a smaller α_1. To see this, assume that country 1 proposes $(\alpha_1 - \varepsilon)$. Rewriting (19) and (18) for $(\alpha_1 - \varepsilon)$, substituting as above and solving yields a share for country 1 equal to φ_1 such that

$$\begin{aligned}\varphi_1 &= \varphi_1 - \frac{(1+\delta)\delta k_1}{(1+k_1)\left(\delta^2+\delta+1\right)}\varepsilon = \varphi_1 - \Theta\varepsilon \\ \text{with } 0 &\leq \Theta \leq 1\end{aligned}$$

Comparing the payoffs for country 1 without and with the deviation, we have

$$\begin{aligned}\Pi_1(\alpha_1, \varphi_1) &= v_1(q^A) + \varphi_1 \Delta - \max(0, k_1(\alpha_1 - \varphi_1)\Delta) \\ \hat{\Pi}_1(\alpha_1, \varphi_1) &= v_1(q^A) + (\varphi_1 - \Theta\varepsilon)\,\Delta - \max(0, k_1((\alpha_1 - \varepsilon) - (\varphi_1 - \Theta\varepsilon))\Delta)\end{aligned}$$

As $\alpha_1 = \varphi_1$ and $(\alpha_1 - \varepsilon) < (\varphi_1 - \Theta\varepsilon) = (\alpha_1 - \Theta\varepsilon)$ because $0 \le \Theta \le 1$, this simplifies to

$$\begin{aligned}\Pi_1(\alpha_1, \varphi_1) &= v_1(q^A, \cdot) + \varphi_1 \Delta \\ \hat{\Pi}_1(\alpha_1, \varphi_1) &= v_1(q^A, \cdot) + (\varphi_1 - \Theta\varepsilon)\,\Delta\end{aligned}$$

and, as $0 \le \Theta$ and $\varepsilon > 0$, we have that $\hat{\Pi}_1(\alpha_1, \varphi_1) - \Pi_1(\alpha_1, \varphi_1) < 0$. Hence, the deviation is not profitable. A similar argument can be developed for the other countries.

The solution shown in the Proposition generalizes the result in (22) for N countries.

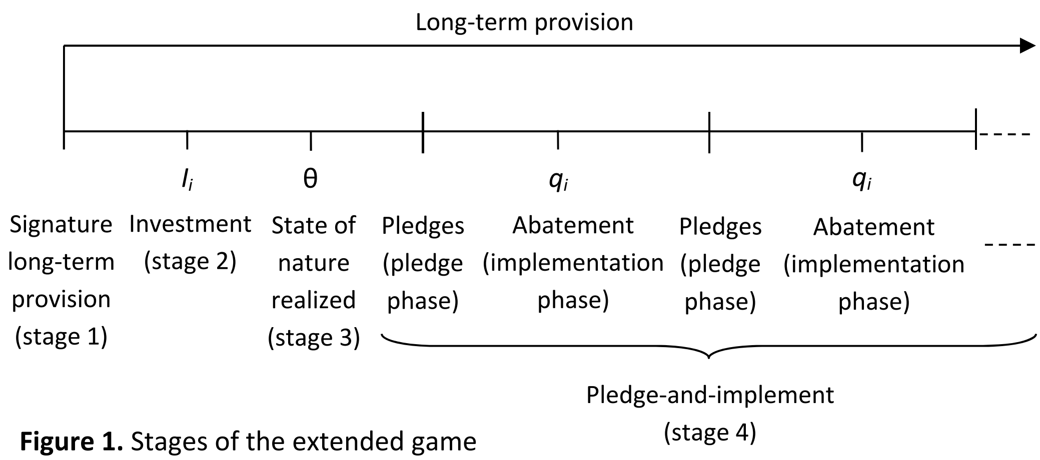

Figure 1. Stages of the extended game

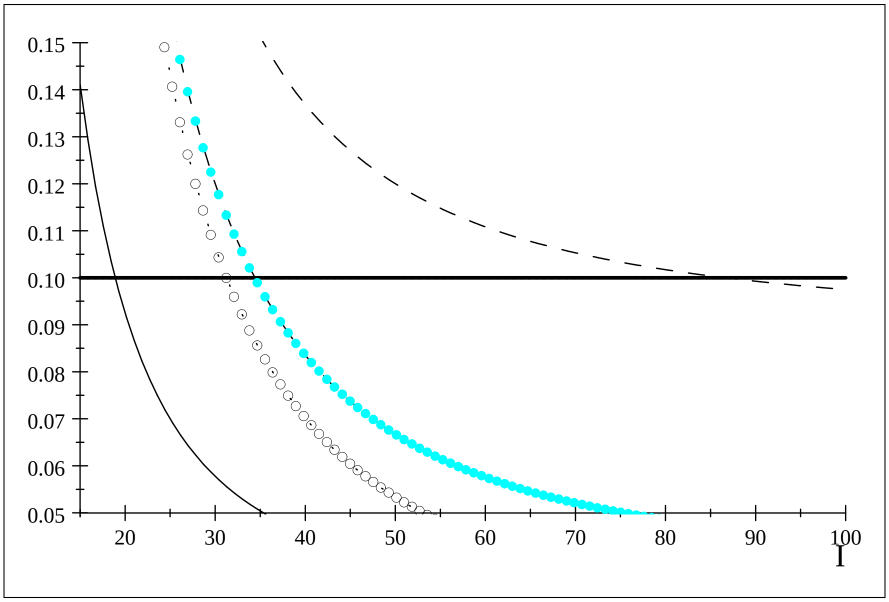

Figure 2. Investment in the first-best and in the "basic Paris Agreement". Thick solid: RHS of (7) and (10). Dashed: LHS of (7). Thin solid: first term on LHS of (10). Full dots: first two terms LHS of (10). Empty dots: LHS of (10).