Formation of Agreements on Climate Change: The Impact of Uncertainty

MICHAEL FINUS
Department of Economics
University of Exeter Business School
Streatham Court, Rennes Drive, Exeter, EX4 4PU
UNITED KINGDOM
m.finus@exeter.ac.uk

PEDRO PINTASSILGO
Faculty of Economics and Research Centre for Spatial and Organizational Dynamics
University of Algarve
Campus de Gambelas, 8005-139 Faro
PORTUGAL
ppintas@ualg.pt

Abstract: This paper addresses the role of uncertainty in the formation of international agreements on climate change. This is undertaken through a coalition formation model. Using different learning scenarios (full, partial and no learning) and different uncertainty cases we show, as opposed to previous studies, that the “veil of uncertainty” is generally not conducive to the success of international agreements. Whenever the uncertainty on the level of the benefits from global abatement is relevant, piercing the “veil of uncertainty”, through more information, leads to more successful agreements.

Key-words: climate change, self-enforcing agreements, transnational cooperation, uncertainty

1. Introduction

Climate change is presently a major challenge to international cooperation, as emphasized by recent studies such as the Stern Review [13] and the IPCC Reports [9]. The establishment of the International Panel on Climate Change (IPCC) in 1988, is usually referred as the first landmark of international cooperation in this domain. The mission of this international body is to provide the most up-date and comprehensive scientific, technical and socio-economic information about climate change [9].

In 1992, the United Nations Framework Convention on Climate Change (UNFCCC) was created in the Earth Summit held in Rio de Janeiro. This treaty aims at stabilizing greenhouse gas concentrations in the atmosphere at a level that would prevent dangerous anthropogenic interference with the climate system [13]. However, the treaty set neither limits to greenhouse gas emissions by individual countries nor enforcement mechanisms.

Only in 1997, a group of 38 countries agreed to specific emission ceilings under the Kyoto Protocol. The target was set in a global emissions reduction of at least 5% below 1990 levels in the commitment period 2008-2012. The protocol entered into force only in 2005 and after several concessions to various participants and the withdraw of the United States of America. Currently, a “Post-Kyoto” agreement for the period after 2012 is being negotiated, with the aim to tighten emission limits, encourage the participation of the US and the “new” emerging polluters China and India.

One of the main problems behind the difficulties in achieving cooperation under international climate agreements is free-riding. Countries have incentives to adopt a non-cooperative behavior: not abating their own emissions and benefiting from the global abatement effort undertaken by the others. In a seminal paper, Hardin [8] shows that this problem is especially virulent for common-pool resources, leading to its over-exploitation. The literature on the formation of self-enforcing international environmental agreements (SEIEAs), based on game theory, also points to the determinant role of free-rider incentives in undermining cooperation (see [5] for a survey).

Uncertainty is also an important element in the formation of international agreements. In fact, there
are large uncertainties on the impact of greenhouse gases on the climate system and on the caused environmental damages. Predictions about abatement costs are also difficult ([9], [13]). For instance, the former US President George Bush used uncertainty as one argument for his decision to withdraw from the Kyoto Protocol. In a letter to Senators, dated March 13, 2001, as quoted by Kolstad [10], he wrote: “I oppose the Kyoto Protocol … we must be very careful not to take actions that could harm consumers. This is especially true given the incomplete state of scientific knowledge”.

A recent strand of literature has approached the formation of international agreements on climate change in the context of uncertainty ([10], [11], [12]). The main conclusion is that under the “veil of uncertainty” the formation of international agreements is more successful. Finus and Pintassilgo ([6], [7]) analyses how general is this conclusion and provide the main driving forces. The present paper summarizes and discusses the authors’ main findings.

2. The Model
Coalition formation is usually modeled using game theory—a branch of mathematics that studies the strategic interaction between decision makers (“players”) by using models usually known as “games”.

2.1 Coalition Formation Game
International environmental agreements are typical single agreements with voluntary participation and open membership, i.e. a country can neither be forced into nor excluded from participation. Therefore, we model coalition formation game as a two-stage open membership single coalition game. In the first stage, players (i.e. countries in our context) decide whether to join an agreement (i.e. a climate treaty in our context) or remain an outsider as a singleton. In the second stage players choose their policy levels (i.e. abatement in our context). The game is solved backward assuming that strategies in each stage must form a Nash equilibrium.

This game has also been called cartel formation game with non-members called fringe players. It originates from the literature in industrial organization (d’Aspremont et al. [1]) and has been applied widely in this literature (see [15] for a survey) but also in the literature on self-enforcing international environmental agreements (see [2] for a survey).

In the first stage, players’ membership decisions lead to a coalition structure, \( K = \{S, I_{n-m}\} \), which is a partition of players, with \( n \) being the total number of players, \( m \) the size of coalition \( S \), \( m \leq n \), and \( N \) the set of players, \( S \subseteq N \). Due to the simple structure of this coalition formation game, that is, there can be at most one non-trivial coalition, coalition structure \( K \) is entirely determined by coalition \( S \). Typically, we will denote a member of \( S \) by \( i \) and call it a signatory and a non-member of \( S \) by \( j \) and call it a non-signatory.

In the second stage, given that some coalition \( S \) has formed, players choose their abatement levels \( q_i \) in our setting. The decision is based on the following payoff function:

\[
\Pi_i = B\left(\sum_{k=1}^{n} q_k\right) - C_i(q_i), \quad i \in N
\]

where \( B(\bullet) \) is country \( i \)’s concave benefit function from global abatement (in the form of reduced damages, e.g. measured against some business-as-usual-scenario) and \( C_i(\bullet) \) its convex abatement cost function from individual abatement. The global public good nature of abatement is captured by the benefit function which depends on the sum of all abatement contributions. For a start, we assume that all functions and their parameters are common knowledge and introduce uncertainty in section 2.3.

Working backward, we assume that the optimal economic strategies in the second stage are the Nash equilibrium of the game between coalition \( S \), with its \( m \) members, and the \( n-m \) singletons. The equilibrium is derived by assuming that coalition members maximize the aggregate payoff of their coalition whereas all singletons maximize their own payoff. The simultaneous solution of these problems leads to the equilibrium abatement levels of signatories \( q'_i(S) \) and of non-signatories \( q'_j(S) \). Their respective equilibrium payoffs in the second stage of the coalition formation game are represented by \( \Pi^*_{i \in S}(S) \) and \( \Pi^*_{j \notin S}(S) \). This assumes no transfers. However, given the assumption of joint welfare maximization of coalition members and the fact that we allow for asymmetric payoff functions, it is perceivable that coalition members share their total payoff \( \Pi_S = \sum_{i \in S} \Pi^*_i(S) \) through transfers \( t_i \) such that the “corrected” payoffs are \( \Pi^*_i(S) + t_i \) with \( \sum_{i \in S} t_i = 0 \).
In the first stage, stable coalitions are determined by invoking the stability concept of internal and external stability, which is de facto a Nash equilibrium in membership strategies. Consider first the version without transfers:

**Internal stability:** \( \Pi^*_i(S) \geq \Pi^*_i(S \setminus \{i\}) \quad \forall i \in S \)

**External stability:** \( \Pi^*_j(S) > \Pi^*_j(S \cup \{j\}) \quad \forall j \notin S \).

That is, no signatory should have an incentive to leave coalition \( S \) to become a non-signatory and no non-signatory should have an incentive to join coalition \( S \). In order to avoid knife-edge cases, we assume that if players are indifferent between joining coalition \( S \) and remaining outside, they will join the agreement. Coalitions which are internally and externally stable are called stable and the set is denoted by \( S^* \). In case there is more than one stable coalition, we apply the Pareto-dominance selection criterion. We denote the set of Pareto-undominated stable coalitions by \( \Psi^* \supseteq S^* \). If non-trivial coalitions are stable, they Pareto-dominate the singleton coalition structure. Note that the coalition structure comprising only singletons is stable by definition and hence existence of an equilibrium is guaranteed.

In the case of transfers, many schemes are perceivable which typically lead to different sets of stable coalitions. In order to avoid this sensitivity, we follow the concept of an almost ideal sharing scheme (AISS) proposed by Eyckmans and Finus [4]. They argue that if and only if:

\[
(2) \sum_{i \in S} \Pi^*_i(S) \geq \sum_{i \in S} \Pi^*_i(S \setminus \{i\})
\]

holds (potential internal stability), then there exists a transfer system which makes \( S \) internally stable. A transfer system for which every potentially internally stable coalition is internally stable belongs to the AISS, which gives each coalition member its free-rider payoff, \( \Pi^*_i(S \setminus \{i\}) \), plus some positive share of the surplus \( \sigma(S) = \sum_{i \in S} \Pi^*_i(S) - \sum_{i \in S} \Pi^*_i(S \setminus \{i\}) \).

Among those coalitions that can be potentially internally stabilized, the AISS stabilizes (in the sense of internal and external stability) those with the highest aggregate welfare over all players. Their conclusion hinges on only one structural property, namely the property of (weakly) positive externalities from coalition formation. It means that whenever a non-signatory joins the climate agreement the other non-signatories become better or at least not worse off. This property (in its strong version) holds in our cartel formation game.

### 2.2 Three Learning Scenarios

We now assume that some parameter values of the payoff functions are uncertain. Following Kolstad and Ulph [11], we assume risk-neutral agents as players are governments and not individuals, and distinguish three learning scenarios: 1) full learning, 2) partial learning and 3) no learning. Full Learning (abbreviated FL) can be considered as a benchmark case in which players learn about the true parameter values before taking the membership decision in the first stage. Hence, uncertainty is fully resolved at the beginning of the game. For Partial Learning (abbreviated PL) it is assumed that players decide about membership under uncertainty but know that they will learn about the true parameter values before deciding upon abatement levels in the second stage. Hence, the membership decision is based on expected payoffs, under the assumption that players will take the correct decision in the second stage. Finally, under No Learning (abbreviated NL) also the abatement decision has to be taken under uncertainty. That is, players derive their abatement strategies by maximizing expected payoffs. The membership decisions are also taken based on expected payoffs, though these payoffs differ from those under partial learning, given that less information is available.

Full learning is certainly an optimistic and no learning a pessimistic benchmark about the role of learning in the context of climate change. Partial learning approximates (because beliefs are not updated in a Bayesian sense) the fact that information becomes available over time. For instance, between the signature of the Kyoto Protocol in 1997, and its entry into force in 2005, with compliance in 2008-12, more information has emerged, as documented by various updated issues of IPCC reports.

### 2.3 Three Uncertainty Cases

We now turn to the assumption about the uncertain parameters of the payoff functions which are summarized in three uncertainty cases. Due to the complexity of coalition formation, the consideration of a particular payoff function, as well as the parameters that are uncertain and their distributions is required. In order to avoid the exclusive focus on the binary equilibrium choices “abate” or “not abate” in the second stage, as for instance in [10] and [11], and to capture the information and the strategic effect, we consider a strictly concave payoff function (used for instance by Barrett [3] and Na and Shin [12] ) which is still simple enough to derive analytical results:
(2) \[ \Pi_i = b_i \sum_{k=1}^{n} q_k - c_i \frac{q_i^2}{2} , \quad i \in N, \quad b_i > 0, \quad c_i > 0 \]

where \( b_i \) is a benefit parameter, \( b_i \sum_{k=1}^{n} q_k \) is the benefit from global abatement, \( c_i \) is a cost parameter, and \( c_i \frac{q_i^2}{2} \) is the abatement cost from individual abatement.

Generally, the benefit as well as the cost parameters could be uncertain. However, in the climate context, uncertainty about the benefits from reduced damages appears to be more important than uncertainty about abatement costs ([10], [11] and [12]). Hence, we simplify the model, by dividing payoffs by the cost parameter \( c_i \), define the benefit-cost ratio by \( \gamma_i = b_i / c_i \), and hence payoff function \( \Pi_i \) reads:

\[ \Pi_i = \gamma_i \sum_{k=1}^{n} q_k - \frac{q_i^2}{2} , \quad i \in N, \quad \gamma_i > 0 . \]

Herein, we call \( \gamma_i \) the benefit parameter. If this parameter is uncertain, then it is represented by the random variable \( \Gamma_i \), with associated distribution \( f_{\Gamma_i} \).

In all three cases, uncertainty is symmetric as all players know as much or little about their own as about their fellow players’ payoff functions. We first lay out the specific assumptions and then provide a wider interpretation.

**Case 1: Uncertainty about the Level of Benefits**

The setting of case 1 is considered in [10] and [11], which the authors call systematic uncertainty as it relates to a common parameter. All players have the same expectations ex-ante, and once uncertainty is resolved, all countries have the same benefit parameter ex-post, which we call symmetric realization, i.e. \( \Gamma_i = \Gamma_k \) \( \forall i, k \in N \). Thus, uncertainty is correlated. However, we find it more illuminating to point out that in this case uncertainty is de facto about the level of the benefits from global abatement. For the latter analysis, it is helpful to point out that this implies that the sum of marginal benefits is uncertain with a positive variance.

Compared to the studies mentioned above, our case 1 appears to be more general in two respects. First, our payoff function does not restrict abatement strategies to a binary choice and hence optimal abatement strategies are a function of the benefit parameter.

Second, we do not assume any particular distribution for the uncertain benefit parameters.

**Case 2: Uncertainty about the Distribution of Benefits**

The setting of case 2 is considered in [12]; uncertainty relates to individual parameters. Though expectations about the benefit parameters are symmetric, their realizations are asymmetric among players. Like [12], we consider that the random variables \( \Gamma_i, \quad \forall i \in N \), are perfectly correlated across all players. Unlike the model of Na and Shin [12] with three players, we consider an arbitrary number of players. Because of the larger complexity, we adopt a specific distribution for parameter \( \Gamma_i \), namely a uniform distribution:

\[ f_{\Gamma_i}(\gamma_i) = \begin{cases} \frac{1}{n} & \text{for } \gamma_i = k, \; k \in N \\ 0 & \text{otherwise} \end{cases} \]

We model perfect correlation by assuming that all players have a different benefit parameter: \( \Gamma_i \neq \Gamma_k, \; \forall i \neq k \in N \). Thus, vector \( \Gamma = (\Gamma_1, ..., \Gamma_n) \) is composed of all the elements of \( N \), i.e. \( \bigcup_{i=1}^{n} \Gamma_i = N \). The sum of marginal benefits is fixed and consequently its variance is zero.

Here perfect correlation implies that uncertainty is purely about the distribution of the benefits from global abatement as the level of global benefits is constant. That is, vector \( \Gamma \) can be viewed as different shares of the global benefits from abatement. Different from [12], we also consider the case of partial learning and most important, the possibility to mitigate asymmetries through transfers, which, as we show later, plays a crucial role for the outcome under full learning.

**Case 3: Uncertainty about the Level and Distribution of Benefits**

Case 3 is a combination of the previous two cases and hence there is uncertainty about common and individual parameters. This translates in our setting into uncertainty about the level and distribution of the benefits from global abatement. This is captured by assuming the same uniform distribution as in case 2, except that all random variables, \( \Gamma_i, \; i = 1, ..., n \), are identically and independently distributed, and therefore uncorrelated. Hence, the sum of marginal benefits is uncertain with positive variance which is larger than in case 2, but smaller than in case 1.
Interpretation of the Three Uncertainty Cases

All three cases capture an important aspect of the uncertainty surrounding climate change. There is much uncertainty about the absolute level of the benefits from reduced damages but also much debate about their regional distribution: which countries will be suffering more from climate change? [14]. Hence, case 3 is the most comprehensive case, but cases 1 and 2 are useful benchmarks in order to isolate effects.

3. Results

In this section the main results of the coalition formation game are presented and discussed.

Proposition I

In uncertainty case 1, under the full, partial, and no learning scenario, expected equilibrium total abatement levels and expected total payoffs are ranked as follows:

1) Total Abatement: \( FL = PL = NL \)
2) Total Payoff: \( FL = PL > NL \)

Thus, if there is only uncertainty about the level of the benefits from global abatement, “learning is good” in terms of payoffs. This result is in stark contrast to [10] and [11].

In case 2, players are ex-ante symmetric though ex-post asymmetric. For partial and no learning this does not affect coalition formation compared to case 1 because players take their membership decisions based on expected payoffs. This does not apply to full learning. Hence, transfers only matter for full learning.

Proposition II

In case 2, under the full, partial and no learning scenario, expected equilibrium total abatement levels and expected total payoffs are ranked as follows:

No Transfers

1) Total Abatement: \( NL = PL > FL \)
2) Total Payoff:

\[
\begin{align*}
NL &= PL > FL & n &= 3 \\
NL &= PL < FL & n &> 3 \\
FL &= PL > NL & n &> 4
\end{align*}
\]

Transfers

1) Total Abatement:

\[
\begin{align*}
FL &= PL = NL & n &< 8 \\
FL &= PL > NL & n &> 8
\end{align*}
\]

2) Total Payoff:

\[
\begin{align*}
FL &= PL > NL & n &= 3 \lor n = 4 \\
FL &= PL > NL & n &> 5
\end{align*}
\]

Thus, we generalize the negative result of Na and Shin [12] about the role of learning by considering more than three players and including the intermediate case of partial learning in the analysis. Even more important, we qualify their conclusion by considering transfers and showing that this conclusion can be reversed, at least for full learning.

Like in case 2, in case 3 players are ex-ante symmetric but ex-post asymmetric. The average degree of asymmetry is positive, therefore larger than in case 1, but smaller than in case 2. Not surprisingly, this improves upon the relative performance of full learning compared to case 2, but weakens it compared to case 1, if there are no transfers. With transfers, like in case 2, heterogeneity becomes an asset under full learning.

Proposition III

In case 3, under the full, partial, and no learning scenario, expected equilibrium total abatement levels and expected total payoffs are ranked as follows:

No Transfers

1) Total Abatement: \( NL = PL > FL \)
2) Total Payoff:

\[
\begin{align*}
FL &= PL = NL & n &= 3 \\
NL &= FL > PL & 4 \leq n &< 8 \\
NL &= FL > NL & n &= 9 \\
FL &= NL > PL & n &> 10
\end{align*}
\]

Transfers

1) Total Abatement:

\[
\begin{align*}
FL &= PL = NL & n &= 3 \lor n = 4 \\
FL &= PL > NL & n &> 5
\end{align*}
\]

Taken together, in case 3 partial learning is always better than no learning and once transfers are introduce full learning ranks first. Only without transfers full learning may rank last but, compared to case 2, this does not happen always, as the degree of asymmetry is
smaller. If we view case 3 as the most relevant case of actual negotiations because it captures uncertainty about the level and the distribution of the gains from cooperation, both relevant in climate change, then our results come to a far less negative conclusion than the previous literature. Even in a strategic context, more information must not necessarily be detrimental to the self-enforcing provision of a public good. However, the larger the uncertainty about the distribution of the gains from cooperation, the more important it is to hedge against free-riding through an appropriate transfer scheme.

4. Conclusion

The results exposed in this paper challenge the conclusion that the “veil of uncertainty” is conducive to the success of international agreements on climate change, as in [10], [11], [12]. This is valid only if there is pure uncertainty about the distribution of the gains from cooperation. However, this is most unlikely in the climate change context. Moreover, should the problem be virulent, it can be mitigated, fixed or even turned into an asset through an appropriate transfer scheme.

References:


