Environmental and Resource Economics

International Environmental Agreements - The Role of Foresight

--Manuscript Draft--

Manuscript Number:	EARE-D-14-00423R3
Full Title:	International Environmental Agreements - The Role of Foresight
Article Type:	Manuscript
Keywords:	Environmental Agreements, Foresight, Stable Set
Corresponding Author:	Eftichios Sophocles Sartzetakis University of Macedonia Thessaloniki, GREECE
Corresponding Author Secondary Information:	
Corresponding Author's Institution:	University of Macedonia
Corresponding Author's Secondary Institution:	
First Author:	Effrosyni Diamantoudi
First Author Secondary Information:	
Order of Authors:	Effrosyni Diamantoudi
	Eftichios Sophocles Sartzetakis
Order of Authors Secondary Information:	
Funding Information:	European Social Fund—ESF (EP-1909-01) — Professor Eftichios Sophocles Sartzetakis
Abstract:	The present paper attempts to bridge the gap between the cooperative and the non-cooperative approach employed to examine the size of stable coalitions, formed to address global environmental problems. We do so by endowing countries with foresightedness, that is, by endogenizing the reaction of the coalition's members to a deviation by one member. We assume that when a country contemplates withdrawing or joining an agreement, it takes into account the reactions of other countries ignited by its own actions. We identify conditions under which there always exists a unique set of farsighted stable IEAs. The new farsighted IEAs can be much larger than those some of the previous models supported but are not always Pareto efficient.
Response to Reviewers:	

International Environmental Agreements - The Role of Foresight *

Effrosyni Diamantoudi
Department of Economics, Concordia University

Eftichios S. Sartzetakis†
Department of Economics, University of Macedonia

February 2017

Abstract

The present paper attempts to bridge the gap between the cooperative and the non-cooperative approach employed to examine the size of stable coalitions, formed to address global environmental problems. We do so by endowing countries with foresightedness, that is, by endogenizing the reaction of the coalition's members to a deviation by one member. We assume that when a country contemplates withdrawing or joining an agreement, it takes into account the reactions of other countries ignited by its own actions. We identify conditions under which there always exists a unique set of farsighted stable IEAs. The new farsighted IEAs can be much larger than those some of the previous models supported but are not always Pareto efficient.

Keywords: Environmental Agreements, Foresight, Stable Set *JEL*: D6, Q5, C7

*The authors are indebted to the Journal's Co- Editor prof. M. Finus and two referees for their insightful comments that greatly improved the exposition of the results. Sartzetakis gratefully acknowledges financial support from the Research Funding Program: "Thalis—Athens University of Economics and Business—Optimal Management of Dynamical Systems of the Economy and the Environment: The Use of Complex Adaptive Systems" cofinanced by the European Union (European Social Fund—ESF) and Greek national funds.

†Corresponding author: Eftichios S. Sartzetakis, Department of Economics, University of Macedonia, 156 Egnatia Street, Thessaloniki 54006, Greece. E-mail: *esartz@uom.edu.gr*

International Environmental Agreements - The Role of Foresight

February 2017

Abstract

The present paper attempts to bridge the gap between the cooperative and the non-cooperative approach employed to examine the size of stable coalitions, formed to address global environmental problems. We do so by endowing countries with foresightedness, that is, by endogenizing the reaction of the coalition's members to a deviation by one member. We assume that when a country contemplates withdrawing or joining an agreement, it takes into account the reactions of other countries ignited by its own actions. We identify conditions under which there always exists a unique set of farsighted stable IEAs. The new farsighted IEAs can be much larger than those some of the previous models supported but are not always Pareto efficient.

Keywords: Environmental Agreements, Foresight, Stable Set *JEL*: D6, Q5, C7

1 Introduction

International cooperation is necessary to address the great number of global scale environmental challenges resulting from the ever increasing human activity, including global warming, ozone depletion, declining fisheries and ocean acidification. Despite the proliferation of global environmental treaties, with more than 120 in force currently, only a handful, such as the Montreal Protocol on Substances that Deplete the Ozone Layer and the North Pacific Fur Seal Treaty, are considered a success (Barrett (2003)). The continuing (from COP-15 in Copenhagen to COP-20 in Lima) stalemate in the negotiations for a post-Kyoto agreement is an illustrative example of the great difficulties in achieving large, stable coalitions to effectively address pressing environmental problems, especially when mitigation costs are significant. Although it is clear that countries will collectively benefit from cooperation, the problem is that each individual country has strong incentives to free ride on the cooperation of the rest. The absence of a global institution empowered with the necessary monitoring and enforcement mechanisms, necessitates the design of self-enforcing international environmental agreements. The environmental economics literature examining the issue of coalition formation using game theory, has followed both the cooperative and the non-cooperative approach, deriving conflicting results concerning the number of countries that will participate in an international environmental agreement. The present paper focuses on bridging the gap between these two main approaches in the theoretical literature.[1]

On the one hand there is a series of works (Carraro and Siniscalco (1993), Barrett (1994), Diamantoudi and Sartzetakis (2006), Rubio and Ulph (2006) and Finus et al. (2006)), that, through a non-cooperative game theoretic approach, supports the pessimistic view that IEAs are futile in the sense that they will be either signed by *very few*[2] countries or, if signed by more[3], it will

[1] For an in depth comparison of the approaches see Tulkens (1998). More general literature reviews are offered by Ioannidis, Papandreou and Sartzetakis (2000) and Finus (2003).

[2] Diamantoudi and Sartzetakis (2006) show that in the model adopted by Barrett (1994) the equilibrium size of an IEA is either 2, 3 or 4 countries. Similarly, Finus and Rundshagen (2001) show that in the model adopted by Carraro and Siniscalco (1993) when transfers are not allowed the equilibrium size of an IEA is 2 or 3 countries.

[3] It is clear, especially from the experience with the Kyoto Protocol, that signing does

be over an agreement without any substance[4]. These results are independent of whether we adopt a simultaneous a la Cournot model or a leadership a la Stackelberg model where the IEA leads and the non-signatories follow. The results of this branch of the literature can be attributed to a fundamental assumption embedded in it, namely, the fact that each country upon withdrawing from the agreement assumes that the agreement will remain intact, at least in terms of membership status. It is not surprising that such a hypothesis, enhances free riding benefits, encouraging deviations and thus, undermines the viability of an agreement.

On the other hand, there is a series of works initiated by Chander and Tulkens (1992) and (1997) that, adopting a cooperative game-theoretic framework, asserts the formation of the grand coalition and the attainment of efficiency. The basic assumption underlying the Chander and Tulkens (1997) model is that when a country withdraws from the agreement it assumes that the agreement collapses and each country fends for itself. Naturally, the pessimistic expectation of a potential perpetrator deters deviations and encourages the sustainment of more cooperative agreements. Their work, however, sheds light to the theory of IEAs from a normative angle: if an agreement is designed in a way that deviators are indirectly yet effectively punished through the collapse of the agreement then cooperation and social optimality may be attainable after all.

The main criticism that the above approach has received concerns the credibility of the threat that the coalition will collapse once a single country deviates and furthermore, whether an agreement based on this assumption is still self-enforcing. The present paper attempts to provide a framework in which countries, when considering deviating from an agreement, behave neither myopically –assuming that the coalition will remain intact– nor phobically –assuming the coalition will completely collapse. Our assumption of farsighted countries provides a more realistic framework, bridging the gap between the two approaches. When a country defects from an agreement it makes no exogenously imposed assumptions regarding the behavior of the

not necessarily imply ratification of an IEA. Although, following the literature we mostly use the term signing (signatory country), our analysis bears on the ratification of an IEA.

[4]Diamantoudi and Sartzetakis (2014) show that using general functional forms the size of the coalition could be significant, but even in such cases the level of welfare does not improve relative to the laissez-faire state where each country optimizes individually.

remaining members of the agreement. Instead, it foresees what their reaction(s) will be, and which equilibrium agreement will result from such an initial deviation. The advantage of farsighted stability is that it considers what happens after an initial deviation in a consistent manner, in accordance with individual optimization behavior. In this paper we examine the case in which farsighted countries always act alone whether they are joining or withdrawing from an agreement. We use the leadership model, assuming that the non-signatories will wait to observe the coalition's decision, behaving thus as followers. Our analysis can easily be applied to the simultaneous model. We find that by not restricting countries to a myopic behavior, increases the set of possible stable coalitions. That is, the assumption of farsighted countries allows for larger coalitions relative to the myopic behavior. Although the grand coalition does not necessarily belong to the set of stable coalitions, it is a possible equilibrium outcome. This result is based on the credible threat of the partial collapse of the agreement once a country withdraws from the agreement. The threat of the partial collapse of the agreement reduces free riding benefits, leading to larger stable coalitions. The main contribution of the present paper is the complete characterization of the farsighted stable set, allowing for the agreement to both shrink and grow in size. That is, we permit renegotiation among countries, in the sense that even if an IEA collapses, countries can always renegotiate a larger agreement. Our analysis provides a solid connection between cooperative and non-cooperative game-theoretic frameworks. Furthermore, we examine the special case of quadratic benefit and damage function to illustrate our results and provide direct comparison with myopic behavior in terms of global emissions and welfare.

Our work relates to the general theoretical framework developed in Ray and Vohra (1997) and (1999) and also to the work by Ray and Vohra (2001) which examines the public good provision problem and focuses on an equilibrium analysis. More recently, Chander (2007) modifies the γ-core concept to reconcile with the coalitional structure in the non-cooperative game theory. In the literature on IEAs, farsightedness has been examined by Ecchia and Mariotti (1998), Carraro and Moriconi (1998) and further developed by Eyckmans (2001) and de Zeeuw (2008). The work by Diamantoudi and Sartzetakis (2015) is closer to the present paper. They examine the case in which any group of countries may choose to coordinate their actions in either joining or withdrawing from an agreement. In contrast, the present paper assumes that countries make their choices independently. Apart from the differences in the theoretical framework, the present paper differs from Diamantoudi and Sartzetakis (2015) in that coordinated action allows countries to use the complete collapse of the agreement as a threat to sustain the grand coalition, which is not the case in the present work.

Our analysis builds upon two critical assumptions, which although common in the IEAs literature, need some justification. First, we examine the equilibrium size of a single coalition, not allowing for multiple coalitions (of size larger than one) to form and co-exist. The literature on IEAs, both the cooperative and the non-cooperative approach, employs this assumption, with very few exceptions, such as Carraro (2003). Although this assumption seems rather restrictive in a general theoretical framework, it is quite realistic in the case of global environmental agreements. All agreements concerning global environmental problems are negotiated and implemented under the United Nations' auspices and they are unique, in the sense that only one agreement addresses each specific problem.

The second restriction we impose is that countries are symmetric. Assuming symmetric countries implies that within each group, signatories and non-signatories, countries emit the same level of the pollutant and enjoy the same level of welfare. Being able to index payoffs in terms of the group rather than to each specific country, allows us to fully characterize the solution set. Furthermore, due to symmetry, each country's payoffs depend only on the size of the coalition and not on the actions of other countries. If we assume heterogeneous countries, the welfare of each country and therefore its decisions will depend on the composition and not only on the size of the coalition, making the identification of stable coalitions extremely difficult. In addition, if countries are assumed heterogeneous, the size of the coalition becomes less important, as the outcome depends mainly on the coalition's membership. For the above reasons, heterogeneity does not allow the derivation of analytical results using generalized models and thus, no significant conclusions and results can be drawn. Although the symmetry assumption is used widely, there are some attempts to introduce heterogeneity. These papers are either based on simulations (Botteon and Carraro (2001) and McGinty (2006)) or linear (Kolstad (2010)) and semi-linear (Fuentes-Albero and Rubio (2010)) models. As to the size of the stable coalition, the results of this literature are not conclusive. Relaxing the symmetry assumption does not always lead to a larger coalition and when it does, aggregate welfare does not necessarily improve.

The rest of the paper is structured as follows. Section 2 describes the model, introducing the necessary notation and some basic results of the myopic behavior literature that are used in the subsequent analysis. In Section 3 the solution concept is modified to capture consistency and foresight while unilateral actions are maintained. Section 4 examines the special case of quadratic benefit and damage functions and Section 5 concludes the paper.

2 The Model

Let $N = \{1, ..., n\}$ denote the set of all countries, each emitting $e_i \geq 0$, $i \in N$, units of a global pollutant as a by-product of the economic activity, with the aggregate emissions denoted by $E = \sum_{i \in N} e_i$. The benefits resulting from the economic activity are expressed as function of country i's emissions, $B(e_i)$, which is assumed to be strictly concave, that is, $B(0) = 0$, $B' \geq 0$ and $B'' < 0$. Each country suffers damages from the aggregate emissions of the global pollutant, $D(E)$, which are assumed to be strictly convex, that is, $D(0) = 0$, $D' \geq 0$ and $D'' > 0$. Thus, country i's welfare function is,

$$w = B(e_i) - D(E).$$

We consider the case that a set of countries $S \subset N$ sign an agreement to reduce the emissions of the global pollutant and $N \backslash S$ do not and we denote the size of coalition by $|S| = s$. Each member of the coalition emits e_s, units of the global pollutant, while each country outside the coalition emits e_{ns}, with aggregate emission level, $E = E_s + E_{ns} = se_s + (n-s)e_{ns}$.

We assume that the coalition acts as a Stackelberg leader. Each country outside the coalition, hereafter called non-signatory country, maximizes its own individual welfare, w_{ns}, knowing the choice of signatories:

$$\omega_{ns}(e_s, s) = \max_{e_{ns}} \left[B(e_{ns}) - D(se_s + (n-s-1)e_i + e_{ns}) \right].$$

The above problem's first order condition,

$$B'(e^*_{ns}(e_s)) = D'(se_s + (n-s)e^*_{ns}(e_s))$$

yields the best response function $e^*_{ns}(e_s, s)$ of the non-signatories. Note that the assumption of symmetry on best response functions has been applied after the optimization.

Members of the coalition, hereafter called signatories, chose their emission's level by maximizing the coalition's welfare, $\sum_{i\in S} w_i = sw_s$, taking explicitly into account the behavior of countries outside the coalition:

$$\omega_s(s) = \frac{1}{s}\max_{e_s}\left[sB\left(e_s\right) - sD\left[se_s + (n-s)e^*_{ns}(e_s, s)\right]\right].$$

The corresponding first order condition,

$$B'(e^*_s(s)) = D'(se^*_s(s)+(n-s)e^*_{ns}(e^*_s(s), s))\left[s + (n-s)\left.\frac{\partial e^*_{ns}(e_s, s))}{\partial e_s}\right|_{e_s=e^*_s(s)}\right]$$

yields the optimal signatories' emissions $e^*_s(s)$.

Diamantoudi and Sartzetakis (2015), provide a proof of the following proposition establishing the properties of the indirect welfare functions of signatory and non-signatory countries, $\omega_s(s)$ and $\omega_{ns}(e^*_s(s), s)$ respectively. Let e_{nc} and E_{nc} denote the individual and aggregate emissions when there is no agreement and countries behave a la Cournot.

Proposition 1 (Diamantoudi and Sartzetakis, (2015)) C*onsider the indirect welfare functions of signatory and non-signatory countries,* $\omega_s(s)$ *and* $\omega_{ns}(e^*_s(s), s)$ *respectively. Let*

$$z^{\min} = \frac{B''(e_{nc}) - nD''(E_{nc})}{B''(e_{nc}) - D''(E_{nc})}$$

then,

1. $e^*_s(s) \gtreqqless e^*_{ns}(s) \Leftrightarrow s \lesseqqgtr z^{\min}$,
2. *if* $s = z^{\min}$ *then* $e^*_s(s) = e^*_{ns}(s) = e_{nc}$,
3. $\omega_s(s)$ *increases (decreases) in* s *if* $s > z^{\min}$ $(s < z^{\min})$,
4. $z^{\min} = \arg\min_{s\in\Re\cap[0,n]}\omega_s(s)$,
5. $\omega_s(s) \gtreqqless \omega_{ns}(e^*_s(s), s) \Leftrightarrow s \lesseqqgtr z^{\min}$[5].

[5]The point of intersection between $\omega_s(s)$ and $\omega_{ns}(s)$ when s is a real number has been identified independently by Rutz and Borek (2000), but the authors did neither identify the relation between the two functions beyond the point of intersection, nor the fact that the point of intersection is the lowest point of $\omega_s(s)$.

The above Proposition establishes that there exists a critical coalition size, determined by adjusting, to the lower integer, the value of, $z^{\min}$, below which signatory countries emit more and attain higher welfare than the non-signatories, and above which the reverse is true. Since $z^{\min}$ denotes the intersection of ω_s with ω_{ns} that lies right at the minimum value of ω_s, it follows that $z^{\min}$ results in the lowest welfare level for each member of the coalition, when compared to other coalition sizes. Furthermore, the above results imply that the grand coalition yields the highest payoff per signatory than any other coalition size, $\omega_s(s) < \omega_s(n)$ for all $s = 1, ..., n-1$. Although we are able to determine that ω_s is monotonically increasing in s, the same is not possible for ω_{ns}. We choose not to force such a property on ω_{ns} in proving the existence of a unique farsighted stable set of farsighted stable coalitions ,in the next Section. In Section 5 we examine the special case of quadratic benefit and damage function, under which both ω_s and ω_{ns} are monotonically increasing in s and we are able to draw welfare conclusions.

In determining the size of the stable coalitions, the main body of the literature (Carraro and Siniscalco (1993), Barrett (1994) and Finus et al. (2006)) adopts the D 'Aspremont et al. (1983) notion of stability. That is, a coalition of size s^* is, internally stable if $\omega_s(s^*) \geq \omega_{ns}(s^* - 1)$ and externally stable if $\omega_{ns}(s^*) \geq \omega_s(s^* + 1)$. According to this, when a country exits (or enters) the coalition, all other agents do not change their behavior in terms of coalition participation adjusting only their emissions as a response to the new size. The general conclusion in this literature is that stable coalitions are small. Diamantoudi and Sartzetakis (2006) show that in the case of quadratic benefit and damage functions the stable coalition is of size not greater than four, that is, $s^* \leq 4$, regardless of n.

3 Unilateral Action under Foresight

Although, it is well established in the literature that the formation of the grand coalition is Pareto efficient (with respect to the agents involved) and that each agent has an incentive to unilaterally free ride on others, the major contribution of D' Aspremont et al. (1983) who first introduced this form of coalitional stability was the observation that if the number of agents is finite, once a member of the coalition deviates and withdraws from the agreement, the remaining coalition will adjust its behavior. If a country withdraws from the agreement its action will be noticed by the coalition and explicitly taken into consideration. Precisely this consideration leads, in our setting, to the adjustment of their emissions in a manner that maximizes the new aggregate welfare of the remaining coalition members.

Once such an adjustment is captured by the model, the result is that it is not always beneficial for a country to withdraw from the agreement. The increase in its welfare the potential deviant may enjoy by increasing its emission level may be offset by the increase in the coalition's emissions as a result of its adjustment.

In the D' Aspremont et al. (1983) model and its variants within the environmental framework it is assumed that, once a country withdraws, the coalition will adjust its emissions. In fact, the coalition's adjustment and the deviating agent's ability to foresee this adjustment is the very merit of the model. It is only natural thus, that we allow the deviating country to *fully* foresee what is going to happen after it withdraws, including changes in other countries' decision to participate in the coalition.

In this work we endow countries with foresight. In particular we build a solution concept in the spirit of von Neumann and Morgenstern (vN-M) (1944) abstract stable set while amending the dominance relation to incorporate forward looking behavior,[6] which gives us a set of stable coalitions that would survive credible deviations.[7] The issue of credibility and foresight has arisen on several occasions in economic models and more fundamentally in solution concepts within an economic or game theoretic context.

In the D' Aspremont et al. (1983) solution concept when an agent contemplates exiting a coalition of size s it compares $\omega_s(s)$, that is, the welfare it enjoys while a member of the coalition, with the welfare it will enjoy once it exits and joins the non-signatories of a coalition of size $s-1$, that is,

[6]A dominance relation is a binary relation that allows us to compare any two elements (in our case IEAs) and select the stable ones. A myopic dominance relation embeds not only the preference by the acting agent but also the ability for that agent to induce the dominant element from the dominated one. A farsighted dominance relation captures a sequence of elements whereas each acting agent prefers the ultimate outcome (end of the sequence) to his status quo, while he can only induce the next step of the sequence.

[7]Diamantoudi and Xue (2007) have applied previously vN-M's stable sets as an equilibrium concept.

$\omega_{ns}(s-1)$. This comparison examines the internal stability condition as defined earlier. The agent implicitly assumes that once it deviates, no one else will want to deviate and therefore it will indeed enjoy welfare $\omega_{ns}(s-1)$ with certainty. But this is not always the case, in fact, it is possible that another country may wish to exit coalition of size $s-1$ by now, and join the non-signatories $s-2$, and so on. Thus, the country should compare its status quo $\omega_s(s)$ with the *final* outcome that will result once it initiates a sequence of events by exiting a coalition of size s. This *final* outcome can be characterized as such only if no more countries wish to exit and no more countries wish to join, if, in other words, it is stable itself. Put differently, we can determine whether a coalition is stable or not, only if we know what every other coalition is. Such a circular approach is adopted by the classical notion of the abstract stable set.

The (abstract) stable set originally defined by von Neumann and Morgenstern (1944) is a solution concept that captures consistency. The stable set approach, instead of characterizing each outcome independently, characterizes a solution set, that is, a collection of outcomes that are stable, while those excluded from the solution set are unstable. Moreover, no inner contradictions are allowed, that is, any outcome in the stable set cannot dominate another outcome also in the stable set[8]. Similarly, every outcome excluded from the stable set is accounted for in a consistent manner by being dominated by some outcome in the stable set[9]. Although the notion of the stable set is very appealing exactly due to the aforementioned properties that attribute consistency[10] it has been criticized on two grounds. Firstly, it does

[8]This feature of the stable set is known as *Internal Stability*, yet we will avoid the terminology since it coincides with the one attributed to coalitions which is entirely different. Characterizing a coalition as internally stable implies that no member wishes to exit the coalition. This latter meaning of internal stability is the one we will maintain throughout the paper as formalized in Definition 1 that follows.

[9]This feature of the stable set is known as *External Stability*. The same problem with terminology arises here as well. We will maintain the meaning of external stability as formalized in Definition 1.

[10]Its appeal is captured and improved upon by Greenberg (1994). In the *Theory of Social Situations (TOSS)*, a unifying approach towards cooperative and non-cooperative game theory, where any behavioral and institutional assumptions are explicitly defined, an equivalence is shown between the von Neumann & Morgenstern (vN-M) stable set and the *Optimistic Stable Standard of Behavior (OSSB)*, a solution concept built in the spirit of vN-M stability, yet with the precise assumption of optimistic behavior explicitly

not always exist, and secondly, it suffers from myopia as well, as depicted by Harsanyi (1974) who suggested that the simple (one step) dominance relation be replaced by indirect dominance that allows agents to consider many steps ahead. His criticism inspired a series of works in abstract environments by Chwe (1994), Mariotti (1997) and Xue (1998) among others. Our work differs from the works of Chwe (1994), Mariotti (1997) and Xue (1998) in two critical ways: (i) they allow for many coalitions to form in equilibrium which in the context of IEAs would imply many parallel and disjoint IEAs and (ii) agents can induce different structures/IEAs by acting in groups and not only unilaterally as in our analysis. Due to the abstract nature of the authors' works only general existence results are obtained. In contrast, we are able to fully characterize the solution set in the context of IEAs.

In the spirit of von Neumann & Morgenstern's (1944) stable set, Harsanyi's (1974) indirect dominance and within the context of coalition formation, we consider a set σ which is the collection of all farsighted stable coalitions. Let C_s denote any coalition of size s, then, $\sigma = \{C_s, C_t..., C_f\}$. A coalition C_s is farsighted stable given σ and thus, $C_s \in \sigma$ if it is both internally and externally farsighted stable given σ.

A coalition is considered to be internally stable if no country wishes to exit from it. So far, a country compared its current welfare $\omega_s(s)$ with the welfare of the set of non-signatories it would join, $\omega_{ns}(s-1)$. We claim that such a comparison is justified only if C_{s-1} is a stable coalition itself, i.e., $C_{s-1} \in \sigma$ as well, which would imply that if C_{s-1} becomes the status quo it would remain so. Otherwise, if $C_{s-1} \notin \sigma$ once at C_{s-1} some other country may wish to exit. Thus, the very first country when contemplating whether to exit from C_s or not it should compare its welfare while a member of C_s to the final stable outcome that will arise. A parallel process describes external stability. A coalition C_s is externally stable if no country wishes to join in. Again the country makes such a decision by comparing its welfare under the status quo $\omega_{ns}(s)$ with the welfare it will enjoy once it joins the coalition, namely $\omega_s(s+1)$. Such a comparison is justified only if $C_{s+1} \in \sigma$

formalized. TOSS amplified the pertinence of stability by recasting the dominance relation into a broader concept beyond the boundaries of a binary relation. In doing so, behavioral assumptions can be imposed on the agents, and more complex institutional settings can be analyzed.

is a stable coalition itself and thus, no more countries wish to enter.[11] The country should compare its status quo with the final outcome that will arise. Formally,

Definition 1 *A set of coalitions, σ, is a farsighted stable set if*

1. *σ is free of inner contradictions:*

   (a) *Every coalition $C_s \in \sigma$ is internally farsighted stable, i.e., there does not exist a finite sequence of coalitions $C_{s-1}, ..., C_{s-m}$, where $m \in \{1, ..., s\}$ such that $C_{s-m} \in \sigma$ and $\omega_s(s-j) < \omega_{ns}(s-m)$ for every $j = 0, 1, ..., m-1$.*

   (b) *Every coalition $C_s \in \sigma$ is externally farsighted stable, i.e., there does not exist a finite sequence of coalitions $C_{s+1}, ..., C_{s+m}$, where $m \in \{1, ..., n-s\}$, such that $C_{s+m} \in \sigma$ and $\omega_s(s+m) > \omega_{ns}(s+j)$ for every $j = 0, 1, ..., m-1$.*

2. *σ accounts for every coalition it excludes:*

   That is, for every coalition $C_s \notin \sigma$ either

   (a) *C_s is internally farsighted unstable, i.e., there exists a finite sequence of coalitions $C_{s-1}, ..., C_{s-m}$, where $m \in \{1, ..., s\}$ such that $C_{s-m} \in \sigma$ and $\omega_s(s-j) < \omega_{ns}(s-m)$ for every $j = 0, 1, ..., m-1$, or*

   (b) *C_s is externally farsighted unstable, i.e., there exist a finite sequence of coalitions $C_{s+1}, ..., C_{s+m}$, where $m \in \{1, ..., n-s\}$, such that $C_{s+m} \in \sigma$ and $\omega_s(s+m) > \omega_{ns}(s+j)$ for every $j = 0, 1, ..., m-1$.*

Note that the null coalition with $s = 0$ that contains no members is trivially internally farsighted stable since there does not exist a country to

[11] Note that through considering decreasing and increasing sequences of coalitions (through internal and external stability) we exhaust all possibilities. Allowing for simultaneous entry and exit does not add to our analysis given that countries are symmetric. Although other definitions of stability could be used, we have chosen the one defined in D' Aspremont et al. (1983), in order to allow for direct comparison with the myopic literature's results.

exit, and that a perfectly collusive situation where the coalition C_n contains all the countries is trivially externally farsighted stable, since there do not exist any more countries to join.

We mentioned earlier that one of the major drawbacks of the stable set is it rarely exists. However, Theorem 1 establishes that there exists a stable set σ.[12].

Another problem associated with the stable set is its multiplicity. More precisely the existence of more than one farsighted stable set σ, suggesting different collections of farsighted stable outcomes. Notice the difference between uniqueness of a farsighted stable coalition and uniqueness of a set of farsighted stable coalitions. The former, is obviously not true as we will illustrate in the following result where we claim that a sequence of coalitions are farsighted stable, whereas the latter is asserted in the following Theorem where we argue that σ is unique.

Theorem 1 *There exists a unique farsighted stable set of farsighted stable coalitions, σ.*

Proof. The proof of Theorem 1 consists of three parts. In the first part we construct set σ with the use of an algorithm that identifies the coalitions. Note that although σ is constructed in terms of coalition sizes for simplicity, it in fact contains all the permutations of each size. They obviously do not dominate each other since when one agreement is induced from another it has to grow or shrink monotonically. In the second part we show that σ is farsighted stable and in the third part we show that σ is the unique farsighted stable set.

The following two graphs illustrate the construction and stability of σ. The coalition size appears on the horizontal axis while the per member payoffs appear on the vertical axis. The solid curve depicts the per member payoffs of signatories, $\omega_s(s)$. Note that the curve, in accordance with the definitions provided in Section 2 and proved in Diamantoudi and Sartzetakis (2015), decreases until $z_{\min}$ and increases thereafter. The dotted curve depicts the per member payoffs of non-signatories, $\omega_{ns}(s)$. Again, the graphs are consistent with Section 2, since $\omega_{ns}(s)$ is non-monotonic and uniquely intersects $\omega_s(s)$

[12]Diamantoudi (2005) within the context of cartel stability in a price leadership model establishes that such a dominance relation is acyclic if the coalition members' payoff is inreasing in the size of the coalition, that is, in our context, when $\omega(s)$ is increasing in s.

from below at $z_{\min}$. Lastly, the dashed curve is $\omega_{ns}(s)$ shifted by one, that is $\omega_{ns}(s-1)$. By shifting $\omega_{ns}(s)$ we can identify the myopic stable coalition size, s^*, clearer: it is the largest integer before the intersection where $\omega_{ns}(s-1)$ cuts $\omega_s(s)$ from below. Note that in Figure 2 we depict the case where s^* is unique, whereas in Figure 1 we depict the case where $\omega_{ns}(s-1)$ intersects $\omega_s(s)$ more than once and hence, each largest integer just below each intersection is a myopically stable coalition.

Construction

Step 1: Let $\mathcal{M}$ be the set of all myopically stable coalitions, that is,

$$\mathcal{M} = \{C_{s^*} \mid \omega_s(s^*) \geq \omega_{ns}(s^* - 1) \text{ and } \omega_{ns}(s^*) \geq \omega_s(s^* + 1)\}.$$

Now construct two sequences of coalitions, $C_{s_1^*}, ..., C_{s_j^*}, ..., C_{s_m^*}$ and $C_{\bar{s}_1}, ..., C_{\bar{s}_j}, ..., C_{\bar{s}_m}$ as follows:

(i) $C_{s_j^*} \in \mathcal{M}$. $C_{\bar{s}_j}$ is any coalition where $0 \leq \bar{s}_j \leq n$ for all $j = 1, ..., m$.

(ii) $C_{s_1^*}$ is the largest coalition in $\mathcal{M}$. $C_{\bar{s}_1}$ is the largest coalition such that $\bar{s}_1 < s_1^*$ and $\omega_{ns}(\bar{s}_1) \geq \omega_s(s_1^*)$.

(iii) $C_{s_j^*}$ is the largest coalition in $\mathcal{M}$ such that $s_j^* \leq \bar{s}_{j-1}$. $C_{\bar{s}_j}$ is such that $\bar{s}_j < s_j^*$ and $\omega_{ns}(\bar{s}_j) \geq \omega_s(s_j^*)$ for all $j = 1, ..., m$.

In Figure 1 there are four myopically stable coalitions corresponding to the four intersections between $\omega_{ns}(s-1)$ and $\omega_s(s)$, that is, $\mathcal{M} = \{s_1^*, s_2^*, s_3^*, s_4^*\}$. These points are marked on the $\omega_s(s)$ curve by a star *. s_1^* is the largest coalition in $\mathcal{M}$. In fact, in this particular case $\mathcal{M}$ coincides with the sequence $C_{s_1^*}, ..., C_{s_j^*}, ..., C_{s_m^*}$. But this does not have to always be the case, $\mathcal{M}$ can strictly contain the sequence $C_{s_j^*}$. In order to identify $\bar{s}_1$ we start from $\omega_s(s_1^*)$ and trace backwards on the $\omega_s(s)$ curve until we find the first (largest) integer for which $\omega_{ns}(s) \geq \omega_s(s_1^*)$. This largest integer becomes $\bar{s}_1$. s_2^* is the largest element of $\mathcal{M}$ such that $s_2^* \leq \bar{s}_1$. Starting from s_2^*, we repeat the same process to identify $\bar{s}_2$, which in this particular case is the same as s_3^*. Finally, starting from s_3^* we find $\bar{s}_3 = s_4^*$. The $\bar{s}_i$ points are marked by a dot on the $\omega_{ns}(s)$ curve. In Figure 2, there is only one myopic coalition, and thus there is no need to define $\bar{s}_j$.

Naturally the $C_{s_j^*}$ sequence stops either when $C_{s_m^*}$ is such that $\omega_s(s_m^*) > \omega_{ns}(s)$ for all $s \leq s_m^* - 1$ or when $C_{s_m^*}$ is the smallest coalition in $\mathcal{M}$. Note that the existence of $C_{s_1^*}$ is established in D 'Aspremont et al. (1983) and reinforced by the results in Diamantoudi and Sartzetakis (2015). The existence of $C_{\bar{s}_1}$, however, is not guaranteed but it is not necessary for the construction of σ as illustrated in Step 2 below.

Figure 1

Step 2: In this step we construct σ. Our stable set has the following form: $\sigma = \{C_{s^{-l}}, ..., C_{s^{-i}}, ..., C_{s^{-1}}, C_{s^0}, C_{s^1}, ..., C_{s^i}, ..., C_{s^k}\}$. Let $C_{s^0} \equiv C_{s_m^*}$. Then C_{s^i} is determined as follows: Consider the smallest s such that $s > s^{i-1}$ and $\omega_s(s) \geq \omega_{ns}(s^{i-1})$. If $s \in \left\{\cup_{j=2}^{m} \left(s_j^*, \bar{s}_{j-1}\right]\right\} \cup (s_1^*, n]$ then let $s^i = s$. Otherwise, if $s \in \left(\bar{s}_j, s_j^*\right]$ for some $j = 1, ..., m-1$ then let $s^i = s_j^*$.

$C_{s^{-i}}$ is the largest coalition such that $s^{-i} < s^{-i+1}$ and $\omega_{ns}(s^{-i}) \geq \omega_s(s^{-i+1})$ for all $i = 1, ..., l$. Note that if C_{s^0} is such that $\omega_s(s^0) > \omega_{ns}(s)$ for all $s \leq s^0 - 1$ then C_{s^0} is the beginning of the sequence and $l = 0$.

In Figure 1, $C_{s^0} \equiv C_{s_4^*}$. Starting from s_4^* and tracing the $\omega_s(s)$ curve upwards, we are looking for the smallest s greater than s^0, for which $\omega_s(s) \geq \omega_{ns}(s^0)$. The integer which is just higher is s^1, and since $s^1 \in (\bar{s}_2, s_2^*]$, then

$s^1 = s_2^*$. Starting from s^1, we repeat the same process to find s^2 which, in this particular case, is equal to $\bar{s}_1$. Through the same process we find s^3, which is equal to s_1^*. In Figure 1, since $C_{s^0} \equiv C_{s_4^*}$, we cannot look for any $C_{s^{-i}}$, and thus, $\sigma = \{C_{s^0}, C_{s^1}, C_{s^2}, C_{s^3}\}$. The stable coalitions are denoted on the $\omega_s(s)$ curve by a circle, $\bigcirc$. In Figure 2, $C_{s^0} \equiv C_{s^*}$. Following the same process as in Figure 1 we find s^1 and s^2. In Figure 2 we have to look to the left of s^0 as well. To do this, we start from $\omega_s(s^0) = \omega_s(s^*)$ and move parallel to the horizontal axis up until we cross $\omega_{ns}(s)$ to a point lying below $\omega_n(s)$. The next integer to this cross point defines s^{-1}, and thus, $\sigma = \{C_{s^{-1}}, C_{s^0}, C_{s^1}, C_{s^2}\}$.

Stability

No inner contradictions: We will show that all the elements inside σ are both internally and externally farsighted stable. If $|\sigma| = 1$, then the one coalition in it is trivially both internally and externally farsighted stable since agents have essentially no other alternative, that is, no other stable coalition toward which they might want to deviate.

If $|\sigma| > 1$, we start by considering $C_{s^{-l}}$. It is internally farsighted stable since no smaller *farsighted stable* coalition belongs to σ, therefore agents have nowhere (farsighted stable) to go to if they exit. Similarly, C_{s^k} is externally farsighted stable since there is no other larger *farsighted stable* coalition.

We start with internal farsighted stability. Let us consider some $C_{s^i} \in \sigma$ where $s^0 < s^i < s^k$. It is internally farsighted stable because $\omega_s(s^i) \geq \omega_{ns}(s^{i-1})$ by construction of σ, therefore, no signatory wishes to exit. Now consider $C_{s^{-i}} \in \sigma$ where $s^{-l} < s^{-i} < s^0$. Observe that if $l > 0$ then for all s such that $s < s^0$ we have $\omega_s(s) > \omega_{ns}(s-1)$. Then, s^{-i} is internally farsighted stable since we cannot construct a sequence that leads from $C_{s^{-i}}$ to $C_{s^{-i-1}}$ as, at least at the last part of the sequence the agent will not wish to exit since $\omega_s(s^{-i-1}+1) > \omega_{ns}(s^{-i-1})$.

We continue with external farsighted stability. Let us consider some C_{s^i} $\in \sigma$ where $s^0 < s^i < s^k$. From the construction of the sequences in Step 1 we can see that if $s \in \left\{\cup_{j=2}^m \left(s_j^*, \bar{s}_{j-1}\right]\right\} \cup (s_1^*, n]$ then $\omega_{ns}(s-1) \geq \omega_s(s)$. Thus, if s^{i+1} is such that $s_j^* < s^{i+1} \leq \bar{s}_{j-1}$ for some j, or $s_1^* < s^{i+1}$ then no sequence of coalitions leading from C_{s^i} to $C_{s^{i+1}}$ can be constructed as, at least at the very last step of the sequence, the agent will not wish to enter since $\omega_{ns}(s^{i+1}-1)$ $\geq \omega_s(s^{i+1})$ and hence C_{s^i} is externally farsighted stable. If $s^{i+1} = s_j^*$ for some $j = 1, ..., m-1$, then again we cannot construct a sequence since $s^i \leq \bar{s}_j < s_j^*$

and $\omega_{ns}(\bar{s}_j) \geq \omega_s(s_j^*)$ and thus s^i is externally farsighted stable. For all s^{-i} where $s^{-l} \leq s^{-i} < s^0$ external farsighted stability comes directly from the construction of σ since $\omega_{ns}(s^{-i-1}) \geq \omega_s(s^{-i})$ and no agents wishes to enter.

Figure 2

Accounting for every exclusion: Consider $C_h \notin \sigma$ such that $h < s^{-l}$. From the construction of σ we have $\omega_s(s^{-l}) > \omega_{ns}(s)$ for all $s < s^{-l}$, hence, C_h is externally farsighted unstable. Consider $C_h \notin \sigma$ such that $s^{-i} < h < s^{-i+1}$ for all $i = 1, ..., l$. From the construction of σ we have that $\omega_{ns}(h) < \omega_s(s^{-i+1})$, hence C_h is externally farsighted unstable.

Now consider $C_h \notin \sigma$ where $s^i < h < s^{i+1}$ for some $i = 0, ..., m$. If $s^{i+1} \in \left\{\cup_{j=2}^{m} \left(s_j^*, \bar{s}_{j-1}\right]\right\} \cup (s_1^*, n]$ then $\omega_s(s^{i+1}) \geq \omega_{ns}(s^i)$ and $\omega_{ns}(s^i) > \omega_s(h)$ for every h such that $s^i \leq h < s^{i+1}$. Thus, C_h is internally farsighted unstable since its members wish to exit and become nonsignatories of C_{s^i}. If $s^{i+1} = s_j^*$ and $h \in (\bar{s}_j, s_j^*)$ for some $j = 1, ..., m-1$ then $\omega_s(s^{i+1}) = \omega_s(s_j^*) > \omega_{ns}(h)$

for every h such that $\bar{s}_j < h < s_j^*$. Therefore, C_h is externally farsighted unstable since nonsignatories of C_h wish to join in and reach $C_{s^{i+1}} \equiv C_{\hat{s}_j}$. If $s^{i+1} = s_j^*$ and $h \in (s^i, \bar{s}_j]$ for $j = 1, ..., m-1$, then $\omega_{ns}(s^i) > \omega_s(h)$ for every h such that $s^i \leq h \leq \bar{s}_j$. Therefore, C_h is internally farsighted unstable since its members wish to exit and become nonsignatories of C_{s^i}.

Lastly, consider the case where $C_h \notin \sigma$ where $s^k < h$ we have $\omega_{ns}(s^k) > \omega_s(h)$ for all $h > s^k$, hence, C_h is internally farsighted unstable.

Uniqueness

Uniqueness stems from the fact that C_{s^0} is always internally and externally farsighted stable, *regardless* of the composition of σ. In particular, it is internally farsighted stable either because $\omega_{ns}(s) < \omega_s(s^0)$ for every $s < s^0$ which means that no agent will ever wish to exit since every situation thereafter is worse than the s^0 *regardless* of σ, or because it is the smallest $s^* \in \mathcal{M}$. In the latter case it is again internally farsighted stable because for all s such that $s \leq s^0$ we have $\omega_s(s) > \omega_{ns}(s-1)$ and no sequence can be constructed that leads to smaller IEAs, *regardless* of σ.

C_{s^0} is externally farsighted stable since, by its very construction we cannot find a sequence that will lead to a larger coalition. Once C_{s^0} is included in every σ two groups of coalitions are excluded from every σ, namely those between C_{s^0} and C_{s^1} due to internal and external farsighted instability and those between C_{s^0} and $C_{s^{-1}}$ due to external farsighted instability. Continuing in this manner we end up with a unique σ as constructed earlier. ■

Therefore, allowing for farsighted behavior increases the set of possible stable coalitions. Within a non-cooperative behavioral framework, the credible threat of the partial collapse of an agreement deters deviations, sustaining thus larger coalitions relative to the case of myopic behavior. The complete characterization of the farsighted stable set confirms earlier indications in the literature, using either simulations or specific functional forms, and bridges the gap between cooperative and non-cooperative game-theoretic frameworks. The following corollary summarizes the relationship between the set of farsighted stable and myopically stable coalitions.

Corollary

1. *There may exist myopically stable coalitions that are not farsighted stable.*

2. *There may exist farsighted stable coalitions that are not myopically stable.*

3. *There may exist farsighted stable coalitions that are smaller than the smallest myopically stable coalition.*

4. *The largest farsighted stable coalition will always be larger than or equal to the largest myopically stable coalition.*

The conclusion we can draw from the full characterization of σ is that, if some international organization proposed any of the farsighted stable IEAs contained in σ, this proposal will be adopted by the countries involved. From a normative point of view when a coordinating agency (such as the UN for example) puts forth a proposal it can always select the one yielding the higher aggregate welfare among the IEAs included in σ. Although we expect that larger coalitions will yield higher welfare, this is only the case for signatories but nothing can be said about the non signatories, given that ω_{ns} is non-monotonic is s. In the next Section we consider the special case of quadratic benefit and damage functions and we are able to draw welfare conclusions. If, however, some agreement that is not farsighted stable is proposed, the final outcome will be a smaller or a larger stable agreement, depending on whether it is farsighted internally or externally unstable respectively.

4 A special case: quadratic benefit and damage functions

In this Section we employ particular functional forms in order to clarify the difference between myopic and farsighted equilibria. In particular, we demonstrate that larger coalitions, and hence higher welfare payoffs, can be attained under foresight even in simple environments with specific functional forms. Since most of the literature is build on the quadratic functional forms paradigm, we use the same functions, that is, $B(e_i) = b\left[ae_i - \frac{1}{2}e_i^2\right]$ and $D(E) = \frac{1}{2}c(E)^2$, where a, b and c are positive parameters.

With these specifications, each country i's welfare function becomes,

$$w(e_i) = b\left[ae_i - \frac{1}{2}e_i^2\right] - \frac{c}{2}\left(\sum_{i\in N} e_i\right)^2 .$$

Solving the corresponding welfare maximization problems, as set up in Section 1, yields the equilibrium values of the emission levels e_s, e_{ns} and E. Substituting these values into the welfare functions, yields the indirect welfare function of the signatories, ω_s, and the non-signatories, ω_{ns},[13]

$$\omega_s = ba^2 \left[\frac{1}{2} - \frac{n^2\gamma}{2\Psi} \right], \text{ and } \quad \omega_{ns} = ba^2 \left[\frac{1}{2} - \frac{n^2\gamma X^2(1+\gamma)}{2\Psi^2} \right] ,$$

where $\gamma = \frac{c}{b}$, $\Psi = X^2 + \gamma s^2$ and $X = 1 + \gamma (n - s)$. The properties of these indirect welfare functions are established in Proposition 1, with $z^{\min} = \frac{1+\gamma n}{1+\gamma}$.

Proposition 1 established the monotonicity of ω_s with respect to s, under general functional forms. The following Lemma shows that ω_{ns} is also monotonically increasing in s, in the special case of quadratic functional forms.

Lemma 1 *Assuming quadratic benefit and damage functions, $\omega_{ns}(s)$ increases (decreases) in s if $s > z_{ns}^{\min}$ ($s < z_{ns}^{\min}$).*

Proof. Consider the derivative of ω_{ns} with respect to s.

$$\frac{\partial \omega_{ns}}{\partial s} = \frac{-a^2 bn^{2\gamma 2}(1+\gamma)\left[1+\gamma\left(n-s\right)\right]\left[(1+\gamma)(s^2\gamma - 2s(1+n\gamma)) + (1+n\gamma)^2\right]}{\left[1+\gamma\left(2n+\left(s-2\right)s+\gamma(n-s)^2\right)\right]^3} .$$

The above expression has the following three roots,

$$n + \frac{1}{\gamma}, \left(n + \frac{1}{\gamma}\right) \left[1 \pm \frac{1}{\sqrt{1+\gamma}}\right] .$$

It is clear that two of the roots, the first and the one with positive sign in the brackets, are of greater value than n, and thus not interesting. The only admissible root $z_{ns}^{\min} = \left(n + \frac{1}{\gamma}\right)\left[1 - \frac{1}{\sqrt{1+\gamma}}\right] < n$, is a minimum, since, evaluated at this point, the second order derivative is positive. ■

Therefore, using quadratic functional forms, both ω_s and ω_{ns} are monotonically increasing in s, the first after $z^{\min}$ and the second after $z_{ns}^{\min}$, intersecting just once at $z^{\min} > z_{ns}^{\min}$. Diamantoudi and Sartzetakis (2006) have determined the size of the unique (myopic) stable IEA, which, as noted above, takes the values of 2, 3, or 4 depending on the values of the parameters.

The unique farsighted stable set of farsighted stable coalitions, σ, is characterized by a sequence whose members satisfy the internal $\omega_s(s^*) \geq$

[13] The solution of the problem is presented in Diamantoudi and Sartzetakis (2006).

$\omega_{ns}(s^* - 1)$ and external $\omega_s(s^*+1) \leq \omega_{ns}(s^*)$ stability conditions. In particular we identify a sequence $s^t = f(s^{t-1})$, for $t = 1, ...$ and $s^0 = \{2, 3, 4\}$ such that,

$$\omega_{ns}(s^{t-1}) = \omega_s(s^t) \Longrightarrow X^2(s^t) + \gamma(s^t)^2 = \frac{\Psi^2(s^{t-1})}{(1+\gamma)X^2(s^{t-1})}.$$

Simple manipulations reduce the above to the following polynomial,

$$\gamma\left(\gamma+1\right)(s^t)^2 - 2\gamma\left(\gamma n+1\right)s^t + (1+\gamma n)^2 - \frac{\Psi^2(s^{t-1})}{X^2(s^{t-1})(1+\gamma)} = 0,$$

Recalling that $s^{\min} = \frac{1+\gamma n}{1+\gamma}$, the roots of the polynomial can be written as follows,

$$(s^t) = s^{\min}\left(1 \pm \frac{1}{\sqrt{\gamma}}\sqrt{\frac{\Psi^2(s^{t-1})}{\left(\gamma n+1\right)^2 X^2(s^{t-1})} - 1}\right).$$

Note that the roots are real, since$\frac{\Psi^2(s^{t-1})}{\left(\gamma n+1\right)^2 X^2(s^{t-1})} - 1 > 0$. This is so, since rearranging terms and selectively substituting the values of X and Ψ yields, $\gamma s\left(s-X\right)\left(\Psi + (1+\gamma n)X\right) > 0$. As shown in Diamantoudi and Sartzetakis (2006), $s > X$ for all $s > s^{\min}$ We can further rule out one of the roots by showing that $1 - \frac{1}{\sqrt{\gamma}}\sqrt{\frac{\Psi^2(s^{t-1})}{\left(\gamma n+1\right)^2 X^2(s^{t-1})} - 1} < 0$.

Thus, the sequence is,

$$(s^t) = s^{\min}\left(1 + \frac{1}{\sqrt{\gamma}}\sqrt{\frac{\Psi^2(s^{t-1})}{\left(\gamma n+1\right)^2 X^2(s^{t-1})} - 1}\right) .$$

It is easy to see that $(s^t) > s^{\min}$ for all t since $\frac{1}{\sqrt{\gamma}}\sqrt{\frac{\Psi^2(s^{t-1})}{\left(\gamma n+1\right)^2 X^2(s^{t-1})} - 1} > 0$ and it can also be easily verified that $\frac{d(s^t)}{d(s^{t-1})} > 0$.

In order to restrict the sequence to integers, let $I\left[x\right]$ denote the integer above x if x is not an integer and $I\left[x\right] = x$ if x is already an integer. Formally the set of farsighted stable coalitions is,

$$\sigma = \left\{ \begin{array}{l} s^t \in [0,n] \mid \\ s^t = I\left[s^{\min}\left(1 + \frac{1}{\sqrt{\gamma}}\sqrt{\frac{\Psi^2(s^{t-1})}{(\gamma n+1)^2 X^2(s^{t-1})} - 1}\right)\right], \\ \text{for } t = 1, 2, ... \text{ and } s^0 = \{2,3,4\} \end{array} \right\}.$$

Given the above set of farsighted stable coalitions, that includes the myopic one, and the result of Remark 1, we are able to draw the following conclusion regarding aggregate welfare.

Remark 1 *Larger farsighted stable coalitions yield higher aggregate welfare, under quadratic specification of benefits and damages.*

To compare myopic and farsighted equilibria, we consider the same numerical example used in Diamantoudi and Sartzetakis (2006), that is, $n = 10$, $a = 10$, $b = 6$ and $c = 0.39999$, which results in $\gamma = 0.066665$. This example yields a myopic stable coalition of $s^* = 3$.

Figure 3 illustrates this numerical example, where $\omega_s(s)$ is depicted by the solid line, $\omega_{ns}(s)$ by the dot-dashed line and $\omega_{ns}(s-1)$ by the dashed line. All three indirect welfare functions are plotted against different coalition sizes s.[14]

Figure 3

Note that coalition $s^* = 3$ is stable since it is both internally, $\omega_s(s^*) > \omega_{ns}(s^* - 1)$ and externally stable, $\omega_s(s^* + 1) < \omega_{ns}(s^*)$. To construct the set of farsighted stable coalitions we follow the process described in the proof of Theorem 1. Starting from the myopic stable coalition $s^0 = s^* = 3$ and tracing the $\omega_s(s)$ curve upwards, we are looking for the smallest s greater than s^0, for which $\omega_s(s) \geq \omega_{ns}(s^0)$. Given that in our example, $\omega_{ns}(3) = -306.321$, the integer which is just higher than s^0, is $s^1 = 5$, and thus $s^1 = s_2^* = 5$.

[14] The three indirect welfare curves in Figure 3 are exactly the same as in Diamantoudi and Sartzetakis (2006).

To find s^2 we repeat the same process, looking for the smallest s for which $\omega_s(s) \geq \omega_{ns}(5) = -19.6664$, which is $s^2 = 9$, and thus $s^2 = s_3^* = 9$. Since in this example $n = 10$, we stop the process and thus, $\sigma = \{3, 5, 9\}$. This exact same process is captured by the sequence σ and it does generate the same numbers.

In terms of aggregate emissions and welfare we note that in this example, the myopic stable coalition yields, $E = E_s + E_{ns} = se_s + (n-s)e_{ns} = 53.3123$ and $W = s\omega_s + (n-s)\omega_{ns} = -3,425.17$. The farsighted stable coalition $s_2^* = 5$, yields, $E = 38.7102$ and $W = -1,501.54$, while farsighted stable coalition $s_3^* = 9$, yields, $E = 16.3158$ and $W = 189.993$. As expected, larger coalitions yield substantially lower aggregate emission levels and higher aggregate welfare levels.

5 Epilogue

We extent the International Environmental Agreements' literature by endowing countries with foresightedness and we derive the size of farsighted stable coalitions. That is, we assume that when a country considers joining or withdrawing an environmental coalition, it takes into account the reaction of all other countries to its own decision. Unlike the case of myopic countries, it does not only consider the other countries emissions' adjustment ignited by its decision, but also the possibility that some of them may change their status relative to the coalition. Thus, in deciding to exit (or join) the coalition, it compares its current welfare, as part of the IEA (or outside the IEA) with its ultimate welfare resulting from its own exit and the reactions by all other countries. We provide a solution concept allowing for unilateral actions that capture exactly such a decision-making process and show that the new farsighted stable coalitions are much larger than those supported when myopic behavior is assumed. Foresightedness introduces the credible threat of the partial collapse of a coalition and in doing so provides a solid connection between the non-cooperative and the cooperative approach. Our results are in accordance with the fact that a number of international environmental agreements with a large number of signatories are in force. Finally, we are able to show that, in the case of quadratic benefit and damage functions, larger farsighted stable coalitions yield lower aggregate emissions and higher aggregate welfare.

Diamantoudi and Sartzetakis (2015) have extended the analysis by considering coordinated actions by group of countries and derived the size of coalitionally farsighted stable IEAs. Our work can further be extended by relaxing the two critical assumptions identified in the Introduction. First, it would be interesting to allow for multiple coalitional structures to form and coexist. Second it is very important to study the case of heterogeneous countries. Although our definition can be trivially extended to accommodate asymmetric decision makers and existence results may be possible to attain, a full characterization of the solution set like the ones offered in this work would be very difficult to obtain.

6 References

1. D' Aspremont, C.A., Jacquemin, J., Gabszewicz, J., and Weymark, J.A. (1983). "On the stability of collusive price leadership." *Canadian Journal of Economics* **16**, 17-25.

2. Barrett, S. (1994). "Self-enforcing international environmental agreements." *Oxford Economic Papers*, **46**, 878-894.

3. Barrett, S. (2003). *Environment & Statecraft: The Strategy of Environmental Treaty-Making*, Oxford University Press, Oxford.

4. Botteon, M., Carraro, C. (2001). "Environmental coalitions with heterogeneous countries: Burden-sharing and carbon leakage." In: Ulph, A. (Ed.), Environmental Policy, International Agreements, and International Trade, Oxford University Press, Oxford.

5. Carraro, C. (2003). "International Environmental Negotiations" in J. Wiener and R. Schriner (eds) International Relations, Encyclopedia of Life Support Systems, Developed under the Auspices of UNESCO, EOLSS Publishers, Oxford, pp. 93-113.

6. Carraro, C. and Siniscalco, D. (1993). "Strategies for the international protection of the environment." *Journal of Public Economics* **52**, 309-328.

7. Carraro, C. and Moriconi, F. (1998). "International Games on Climate Change Control." FEEM working paper, 56.98.

8. Chander, P. (2007). "The gamma-core and Coalition Formation." *International Journal of Game Theory* **35**, 539-556.

9. Chander, P. and H. Tulkens (1992). "Theoretical Foundations of Negotiations and Cost-sharing in Transfrontier Pollution Problems." *European Economic Review* **36**, 388-398.

10. Chander, P. and H. Tulkens (1997). "The Core of an Economy with Multilateral Environmental Externalities." *International Journal of Game Theory* **26**, 379-401.

11. Chwe, M. S.-Y. (1994). "Farsighted Coalitional Stability." *Journal of Economic Theory* **63**, 299-325.

12. de Zeeuw A. (2008). "Dynamic effects on the stability of international environmental agreements." *Journal of Environmental Economics and Management* **55**, 163-174.

13. Diamantoudi, E. (2005). "Stable Cartels Revisited." *Economic Theory* **26**, 907-921.

14. Diamantoudi, E. and Xue, L. (2007). "Coalitions, agreements and efficiency." *Journal of Economic Theory* **136**: 105-125.

15. Diamantoudi, E. and Sartzetakis, E. (2006). "Stable International Environmental Agreements: An Analytical Approach." *Journal of Public Economic Theory **8**, 247–263.*

16. Diamantoudi, E. and Sartzetakis, E. (2015). "International Environmental Agreements - Coordinated action under foresight." *Economic Theory,* **59**, 527–546.

17. Ecchia, G. and Mariotti, M. (1998). "Coalition Formation in International Environmental Agreements and the Role of Institutions." *European Economic Review* **42**, 573-582.

18. Eyckmans, J. (2001). “On the Farsighted Stability of the Kyoto Protocol.” Working Paper Series, Faculty of Economics and Applied Economic Sciences, University of Leuven, No. 2001-03

19. Finus, M., Van E. Ierland, and R. Dellink (2006). ”Stability of climate coalitions in a cartel formation game.” *Economics of Governance* **7**, 271–291.

20. Finus, M. (2003). “Stability and design of international environmental agreements: the case of transboundary pollution.” in H. Folmer and T. Tietetenberg (eds) The International Yearbook of Environmental and Resource Economics 2003-04, Edward Elgar, Cheltenham, pp 82-158.

21. Finus, M. and Rundshagen, B. (2001). “Endogenous Coalition Formation in Global Pollution Control.” FEEM, working paper, 43.2001.

22. Fuentes-Albero, C., Rubio, S.J. (2010) ”Can international environmental cooperation be bought?” *European Journal of Operational Research* **202**, 255–264.

23. Greenberg, J. (1990). *The Theory of Social Situations: An Alternative Game-Theoretic Approach.* Cambridge University Press.

24. Harsanyi, J. C. (1974). “Interpretation of Stable Sets and a Proposed Alternative Definition.” *Management Science* **20**, 1472-1495.

25. Ioannidis, A., Papandreou, A. and Sartzetakis E. (2000). “International Environmental Agreements: A Literature Review.” GREEN working paper, Universite Laval 00-08.

26. Kolstad, C. D. (2010). ”Equity, Heterogeneity and International Environmental Agreements.” The B.E. Journal of Economic Analysis & Policy, **10** (2), Article 3.

27. Mariotti, M. (1997). “A Model of Agreements in Strategic Form Games.” *Journal of Economic Theory* **74**, 196-217.

28. McGinty, M., (2007). ”International environmental agreements among asymmetric nations.” *Oxford Economic Papers* **59**, 45–62.

29. Ray, D. and Vohra, R. (1997). "Equilibrium Binding Agreements." *Journal of Economic Theory* **73**, 30-78.

30. Ray, D. and Vohra, R. (1999). "A Theory of Endogenous Coalition Structures." *Games and Economic Behavior* **26**, 286-336.

31. Ray, D. and Vohra, R. (2001). "Coalitional Power and Public Goods." *Journal of Political Economy* **109***,* 1355-1384.

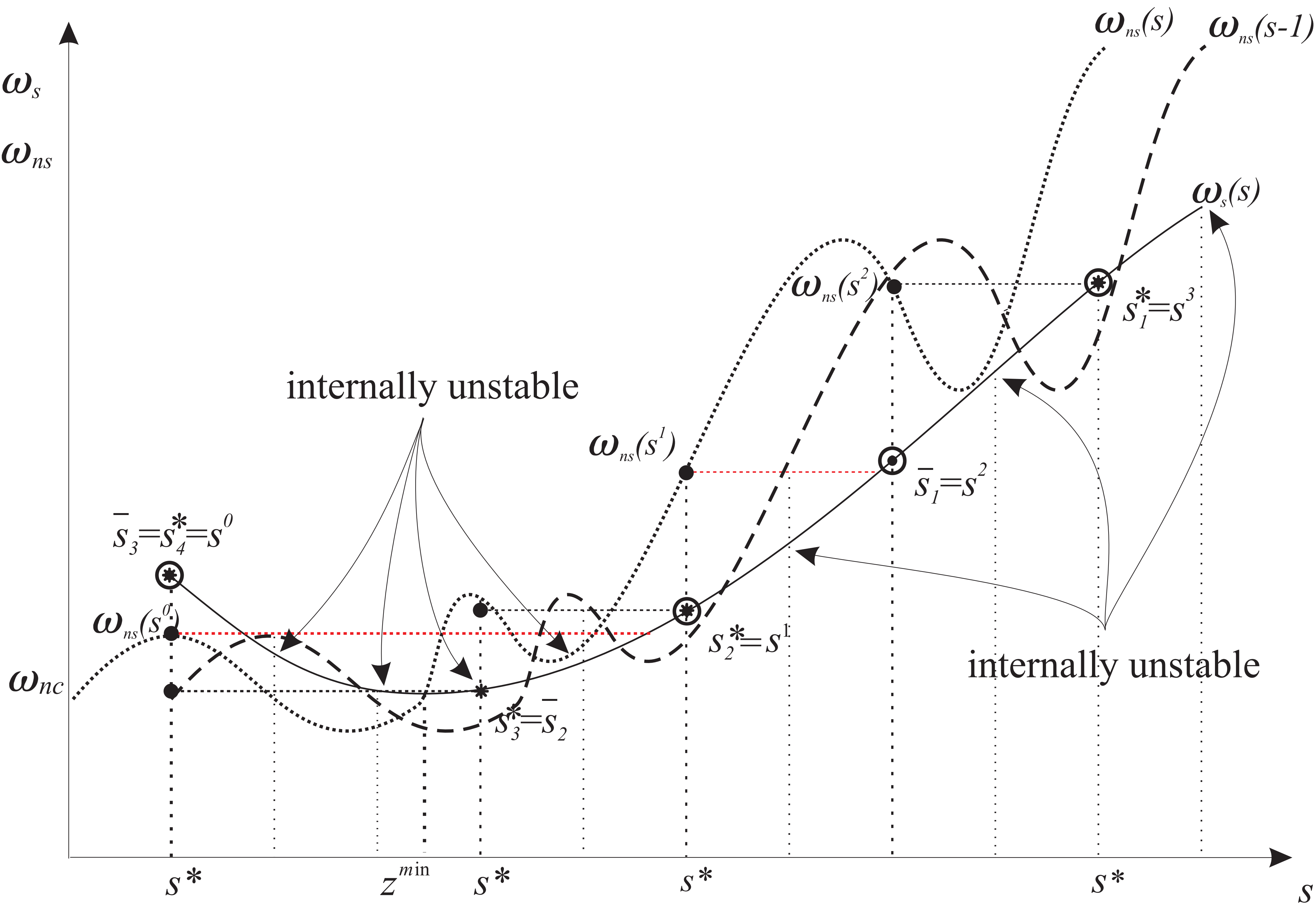
32. Rubio, J. S., and U. Ulph (2006). "Self-enforcing international environmental agreements revisited." *Oxford Economic Papers* **58**, 223–263.

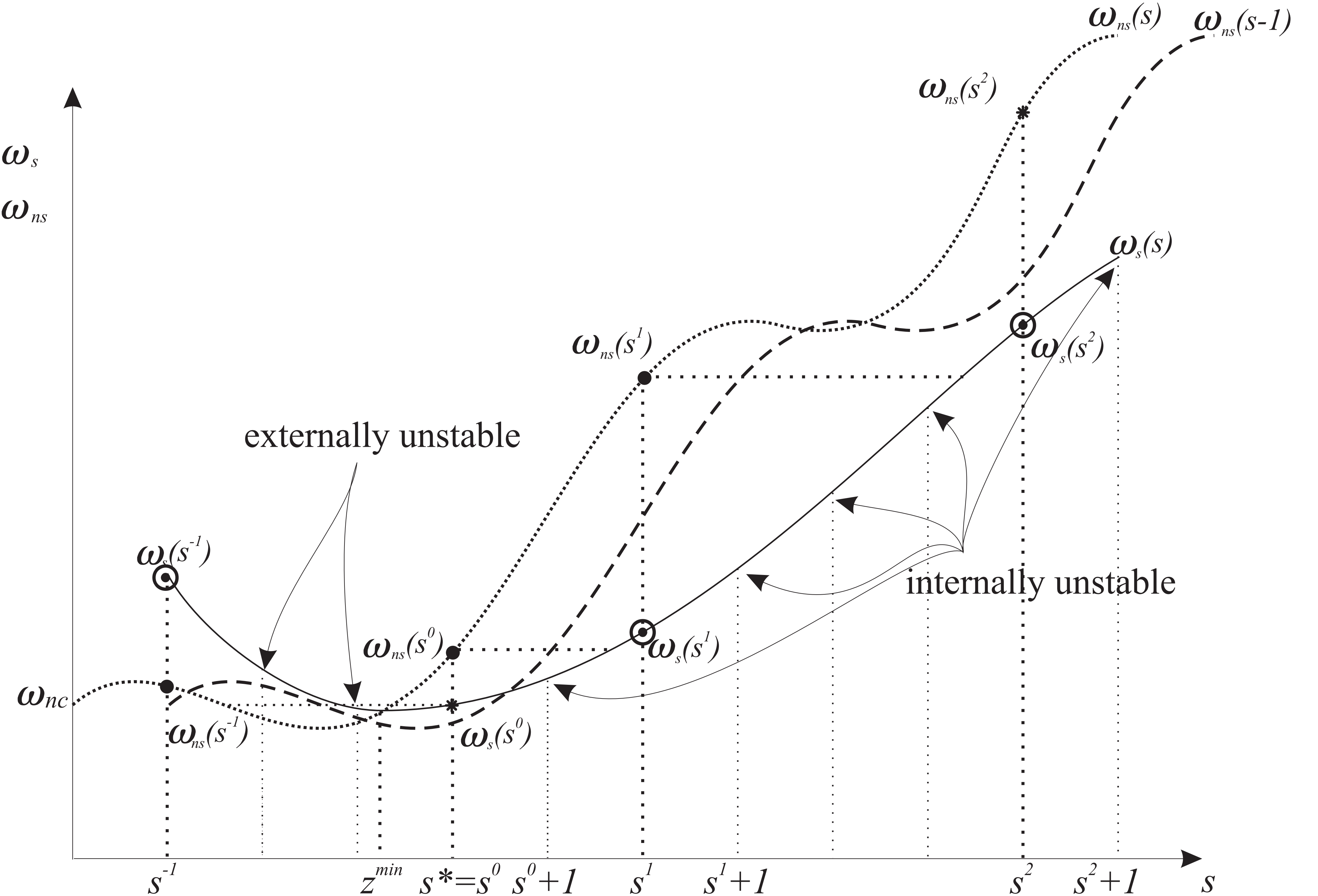
33. Tulkens, H. (1998). "Cooperation versus Free-Riding in International Environmental Affairs: Two Approaches." In Nick Hanley and Henk Folmer (editors): *Game Theory and the Environment*, E. Elgar, Cheltenham, UK.

34. von Neumann, J. and Morgenstern, O. (1944). *Theory of Games and Economic Behavior*. Princeton University Press.

35. Xue, L. (1998). "Coalitional Stability under Perfect Foresight." *Economic Theory* ***11*** *603-627.*

line figure Click here to download line figure Figure1.eps




$\omega_{ns}(s)$
$\omega_{ns}(s-1)$
$\omega_{ns}(s^2)$
ω_s
ω_{ns}
$\omega_s(s)$
$\omega_s(s^2)$
$\omega_{ns}(s^1)$
externally unstable
internally unstable
$\omega_s(s^{-1})$
$\omega_{ns}(s^0)$
$\omega_s(s^1)$
ω_{nc}
$\omega_{ns}(s^{-1})$
$\omega_s(s^0)$
s^{-1}
z^{min}
$s^*=s^0$
s^0+1
s^1
s^1+1
s^2
s^2+1
s

colour figure

Click here to download line figure farshighted simulations figure.jpg