FREE-RIDING AND COOPERATION IN ENVIRONMENTAL GAMES

by

Ana Espínola-Arredondo

M.Sc. in Economics, Universitat de Barcelona, 2002

B.S. Ingeniero Comercial, Universidad de Valparaíso, 1998.

B.A. Business Administration, Universidad de Valparaíso, 1996

Submitted to the Graduate Faculty of

the School of Arts and Sciences in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

University of Pittsburgh

2008

UNIVERSITY OF PITTSBURGH

SCHOOL OF ARTS AND SCIENCES

This dissertation was presented

by

Ana Espínola-Arredondo

It was defended on

March 28th, 2008

and approved by

Andreas Blume, Department of Economics, University of Pittsburgh

Gene Gruver, Department of Economics, University of Pittsburgh

Utku, Unver, Department of Economics, University of Pittsburgh

Esther Gal-Or, Professor, Katz Graduate School of Business, University of Pittsburgh

Dissertation Director: Andreas Blume, Department of Economics, University of Pittsburgh

FREE-RIDING AND COOPERATION IN ENVIRONMENTAL GAMES

Ana Espínola-Arredondo, PhD

University of Pittsburgh, 2008

The dissertation examines free-riding behavior and externality problems using game theory and mechanism design. Specifically, I study free-riding behavior in the negotiation process of International Environmental Agreements. I analyze how countries' noncompliance in an environmental agreement affects the results of the bargaining stage. This study explains why countries fulfill non-enforceable treaties and why some countries want to specify high commitment levels to other countries if there is no international organization that perfectly enforces the contents of the environmental agreement.

The second part of the thesis studies governments' conservation programs. I assume that the production of biodiversity from these programs can generate negative or positive externalities on those nonparticipating landholders. I identify what the government's optimal transfer is when externalities are considered. Finally, the third chapter analyses an alternative definition of kindness. Specifically, we consider that a player (follower) is concerned about those actions that the other player (leader) does not choose. We show that, without relying on interpersonal payoff comparisons (i.e., with strictly individualistic preferences), our model predicts higher cooperation among the players than standard game-theoretic models.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

PREFACE

I am especially grateful to Andreas Blume, he has strongly encouraged me to pursuit my research and gave me the tools to analyze these particular issues in environmental economics. I owe a great debt of gratitude to Esther Gal-Or, Gene Gruver, Santiago Rubio and Utku Ünver for all their helpful comments. I specially would like to thanks my husband, Felix Munoz-Garcia, he has been my main support and motivation during all these years. We have grown together during all these academic process and we are certainly ready to keep dreaming together. I would also like to thank all participants of the 7th Meeting on Game Theory and Practice Dedicated to Energy, Environment and Natural Resources at HEC Montréal, the Public Economic Theory Conference at Vanderbilt University, the 8th BIOECON Conference on "Economic Analysis of Ecology and Biodiversity" at University of Cambridge and the Environmental and Resource Economics Workshop at the University of Colorado at Boulder for their insightful comments. Finally, I would like to thank my father for his constant advice and for being my inspiration as a teacher. My brother Luis who has been a model of ethical values. My officemates Sunita and Jared who have shared their friendship and support in a very crucial time of my life. I am extremely blessed of having them as classmates. All the wonderful people I met in Pittsburgh Ann, Iñaki, Mamen, Lucrecia, Vincent, Marco, Merce, Josep and Anna, they have taught me the meaning of friendship. Thanks are also due to the Department of Economics of the University of Pittsburgh, all my research will not be possible without its trust and support.

1.0 INTRODUCTION

This dissertation uses microeconomic theory to understand environmental problems. Specifically, I study how a local or international authority can solve free-riding problems when it has to allocate a national or transnational public good.

The first chapter, "Free-riding and cooperation in environmental games" examines the negotiation of an international environmental agreement in which different countries determine the (non-enforceable) promises of investment in clean technologies to be included in the agreement. Furthermore, it analyzes countries' optimal investment in emission-reducing technologies, considering that, in addition to the utility that a country perceives from an improved environmental quality, it is also concerned about the relative fulfillment of the terms specified in the international agreement either by itself or by others. I show, first, why countries may prefer to shift most promises of investment in clean technologies to other countries, despite the fact that these promises are usually non-enforceable by any international organization. Second, I determine countries' optimal investments in these technologies, and analyze how their particular investments depend on how demanding the international agreement is, and on the importance that countries assign to each others' relative fulfillment of their part of the treaty. The contribution of this chapter to the literature on environmental games is then twofold. First, it endogenizes the particular commitments that countries include and sign in the IEA, explaining also why countries care about the specific terms of the agreement despite these terms being non-enforceable. In contrast, standard environmental games assume that countries simply decide whether to participate in the IEA. This makes the specific commitments included in the IEA exogenous to the game. Second, this study explains the interaction between the specific terms of an agreement and countries' relative fulfillment of such agreement, which are considered to be completely independent in the existing literature.

The second chapter, "Green auctions: a biodiversity study of mechanism design with externalities" studies the optimal transfer that the government pays when taking into account the externality (either negative or positive) created by a landholder who produces biodiversity. In particular, this paper uses a mechanism design approach to study the biodiversity improvement in a territory, where the government is the principal and the landholders are the agents. I analyze an optimal mechanism that considers multidimensional messages which includes both the biodiversity improvement of the project and its cost. Additionally, this mechanism incorporates the externality (either positive or negative) that a biodiversity project causes in the surrounding agents who decided not to participate. Specifically, I assume that externalities enter in the cost function of the nonparticipating landholders. I show that, in the case of negative externalities, the government will implement a transfer function which is decreasing in the landholder's efficiency level. On the other hand, in the case of a positive externality, paradoxically the government may be interested in the nonparticipation of the most efficient landholders.

Finally, the last chapter "The importance of Foregone Options: generalizing social comparisons in sequential-move games" is a joint work with Felix Munoz-Garcia. In this study we recognize the fact that recent experimental evidence supports the influence of a player's unchosen alternatives on other agent's actions. This study examines a tractable theoretical model of reference-dependent preferences in which individuals compare other players' chosen actions with respect to their unchosen alternatives. We analyze the equilibrium prediction in complete information sequential-move games, and compare it with that of standard games where players are not concerned about unchosen alternatives. We show that, without relying on interpersonal payoff comparisons (i.e., with strictly individualistic preferences), our model predicts higher cooperation among the players than standard game-theoretic models. In addition, our framework embodies different behavioral models, such as social status acquisition and intentions-based reciprocity, as special cases. Finally, we confirm our results in three economic applications: the ultimatum bargaining game, the labor market gift exchange game, and the sequential public good game.

2.0 FREE-RIDING & COOPERATION IN ENVIRONMENTAL GAMES

2.1 INTRODUCTION

Multiple international environmental agreements have been implemented in recent years trying to achieve greater cooperation among countries in their reduction of greenhouse gases, emissions leading to ozone layer depletion, and many other pollutants. For instance, the Montreal protocol (1987) and the Kyoto protocol (1997) establish standards for reductions in the emission (and production) of these environmental damaging products and by-products. Most of these agreements, however, have been very asymmetrically implemented by those countries signing them.

Different economic models have been used to analyze countries' behavior towards such international environmental agreements (IEA henceforth), and especially to analyze why they decide not to carry out the reduction in emissions they sign in these treaties. In particular, most of them deal with IEA as a standard public good game, in which countries incur a (private) cost in reducing emissions in their own country, but benefit from a (public) improved global environmental quality. Since, in addition, the private costs from reducing emissions are usually assumed to be higher than the per country benefits of improved world environmental standards, the amount of pollution that every country decides to reduce in the Nash equilibrium of the game is clearly below the Pareto efficient level. Hence, the individual incentives of every country to free-ride on the environmental quality that other countries provide leads to an under-provision of improved environmental standards.

Despite the fact that the equilibrium resulting from these models predicts the commonly observed practice of free-riding in environmental games, there are some observed behaviors that are difficult to rationalize. First, why do countries want to impose high commitment levels on other signatories if there is no international organization that perfectly enforces the content of the IEA? And second, why do certain countries respect the agreements they acquire in IEAs to a great extent (in spite of their non-enforceability), while others do not fulfill their agreements? This paper proposes a model that supports, first, the interest of a country in imposing high demands on other countries –in terms of the reduction of emissions the IEA specifies for them– during the negotiation stage of the IEA, in spite of the non-enforceability of such agreements. In addition, it explains why certain countries may prefer to invest in emission-reducing technologies even when other countries do not invest, and how this optimal investment depends on how demanding (or conservative) the goals of the IEA are, among other parameters.

Similar to standard public good games, this study considers that every country benefits from the global environmental quality achieved by the overall reduction in emissions, and it incurs a private cost in doing so. In addition, the paper assumes that countries benefit from the relative fulfillment of the agreement (i.e., the extent to which the goal of the IEA is fulfilled). Specifically, I consider that countries can benefit from the relative fulfillment of the IEA because of their *own* and/or because of *other countries'* relative fulfillment of the agreement.[1] In the first case, countries benefit from respecting the terms of the IEA since deviating from their environmental commitments may be severely punished by environmentally oriented citizens ("green voters"), whereas sticking to the terms of the agreement may be rewarded by these voters' support in future elections. In the second case, in contrast, countries benefit from observing that other cosigners fully carry out their promises –i.e., they infer a strong commitment with the fulfillment of the environmental standards included in the IEA— and experience disutility from such lack of commitment otherwise.

[1] Both assumptions can simultaneously be introduced in the model. However, this generalization reduces the intuition of the results without improving its explanatory power.

Different real life observations support the idea that countries may benefit from the relative fulfillment of the agreements in which they participate. First, regarding the positive relationship between voters and countries' relative fulfillment of its commitments, there is strong empirical evidence suggesting that voters do vote for an incumbent politician based on her past performance (relative to her initial promises), what is referred as "retrospective voting."[2] Moreover, in the specific relationship between green voters and countries' relative fulfillment of the IEA, table 1 (appendix) shows the existence of a positive correlation between the proportion of green parties in a country's parliament and that country's relative fulfillment of the commitments it signed in the Kyoto protocol. Second, regarding countries' concern for each others' fulfillment of the IEA, we can also find many real cases, where for instance, certain northern European countries such as Germany –which essentially stick to the terms of the Kyoto protocol– may feel some disappointment from observing that many other signatories do not carry out their promises as they should.[3]

In order to examine the role of these non-binding IEAs in countries' environmental policies, this paper analyzes a two-stage complete information environmental game. Countries first decide the environmental goals to be included in the IEA (negotiation game), and given these goals, they simultaneously choose how much to invest in emission-reducing technologies during the second stage of the game (investment game). In particular, this study shows that countries try to impose the most demanding environmental standards on other countries but not on themselves during the negotiation stage of the game. Indeed, when countries are concerned about either their own relative fulfillment of the IEA or about other countries' relative fulfillment, the specific commitment they sign in the treaty becomes relevant, even if these commitments are non-enforceable by any international organization. These predictions are confirmed by the effort that countries exert trying to achieve that other countries sign different non-binding international agreements, either on environmental issues or not.[4] In

[2] For example, Francis *et al.* (1994) find that representatives whose voting records are closer to the predicted senatorial position for their state are more likely to enter a primary, which supports the hypothesis that representatives expect primary voters to choose retrospectively.

[3] Existing evidence suggests that only 15 out of the 41 countries included in Table 3 (appendix) of the Kyoto protocol have fulfilled their commitments in Article 3 (which specifies a 5 percent reduction in the emission of greenhouse gases from 1990 to 2008).

[4] For instance, in February 2007 the European Union insisted that they would sign a reduction in emissions

addition, the research identifies how the commitments involved in the IEA affect countries' investment in clean technologies in a second stage of the game, despite of the fact that these commitments are non-binding. Specifically, it determines under what parameter values (and under what commitments signed during the negotiation stage) countries' reduction of pollutants is higher than what standard environmental games predict. Finally, the paper explains the case where a country would only accept an agreed level equal to zero (e.g., the case of United Kingdom in the Helsinki protocol or U.S. in the Kyoto Protocol), however it invests positive amounts in clean technologies in the investment game.

The contribution of this paper to the literature on environmental games is then twofold. First, it *endogenizes* the particular commitments that countries include and sign in the IEA, explaining also why countries care about the specific terms of the agreement despite these terms being non-enforceable. In contrast, standard environmental games assume that countries simply decide whether to participate in the IEA, whose environmental goals coincide with the Pareto-efficient level (probably determined by scientists). This makes the specific commitments included in the IEA *exogenous* to the game. Second, this paper explains the interaction between the specific terms of an agreement and countries' relative fulfillment of such agreement, which are considered to be completely independent in the existing literature. Interestingly, this model can be applied to many other settings, where players (either countries, firms, or individuals) interact with other players signing a contract in which they both engage in the provision of a certain public good. The contents of the contract are observable, but cannot be enforced by a third party, such as a court of law. Specifically, if players are concerned about the relative fulfillment of the contract by other players, or by themselves, this model predicts higher contribution levels, and lower free-riding behaviors than in standard public good games.

The paper is organized as follows. Section two comments on the game-theoretic literature that studies IEAs. In particular, it focuses on those analyses predicting higher levels of

by 30 percent if other heavy pollutants (e.g. U.S., China and India) sign the agreement as well. Another example is the Byrd-Hagel Resolution, passed by the U.S. Senate on July 25, 1997, which stated that the United States should not be a signatory to any protocol that did not include binding targets and timetables for developing as well as industrialized nations.

emission reduction than in standard environmental games. Section three describes the model under the assumption that countries are concerned about their own relative fulfillment of the agreement (e.g., because of the importance of green voters), and analyzes countries' best response functions in this setting. Afterwards, it examines countries' equilibrium strategies in this simultaneous move game and analyzes the optimal proposals to be made by every country during the (previous) negotiation stage of the IEA, given the above equilibrium strategies. Section four describes the model under the assumption that countries care about each others' relative fulfillment of the IEA and compares its results with those of the first model. Finally, the last section elaborates on the conclusions of the paper, as well as further areas of research.

2.2 RELATED LITERATURE

Given the relatively pessimistic prediction of the existing literature examining environmental improvement as a (global) public good, many different game-theoretic approaches have been applied to explain why cooperation is sometimes observed in international environmental policies. In the following subsection, I elaborate on the main results and assumptions of these models –grouped into four different branches– as well as their main criticisms. Additionally, subsection 2.2 summarizes the debate on treaty compliance.

2.2.1 Literature on environmental games predicting cooperation

In recent years, different authors have used the theory of repeated games to rationalize why certain international agreements are in fact respected along time, see Barrett (1994a, 1994b, 1999), Cesar (1994) and Rubio and Ulph (2007). In particular, a cooperative solution can be supported as a Nash equilibrium of the repeated game when countries' discount factor is high enough. Despite their satisfactory results in terms of cooperation among the players, the disadvantage of using repeated games to analyze such interactions among countries implementing IEAs are, among others: (1) the restriction on the sufficiently high values for the discount factor, which is difficult to reconcile with myopic politicians, and (2) the multiplicity of equilibria supported as the Nash equilibrium of the repeated game, which limits the predictive power of the model.

Similarly, another class of models considers countries' preferences for "international equality", as in game-theoretic models analyzing social preferences with inequity adverse agents, see Hoel and Schneider (1995). In this case, the equilibrium prediction also determines that countries fulfill the agreement they sign in the IEA, at least to a greater extent than in the standard models described above. That is, their reduction of pollutant emissions is more relevant when countries have social preferences among other countries than when they only have strictly individualistic preferences. Also, Jeppesen and Anderson (1998) develop Barret's model (1994a) incorporating the idea of fairness introduced by Rabin (1993). They show that if countries are highly concerned about the welfare of other countries, full cooperation can be supported as an equilibrium. However, the assumption that countries are actually concerned about the payoffs that another country obtains from playing this environmental game does not seem to be very realistic.

Another class of models explaining the seeming dissonance between the standard theoretical models analyzed above –in which countries are predicted to have poor incentives to reduce emissions— and real cases –in which certain countries carry out their promises in the IEA to a great extent— uses the possibility that an international organization imposes sanctions on the "defecting" countries, see Barrett (1992, 1994a).[5] Obviously, introducing the possibility of receiving a sanction induces every country to maintain its promises in the IEA. However, these models have also been subject to criticism in the literature, since they assume enforceable contracts. Given that most of these international agreements cannot be enforced by any legal organization, this model is probably very restrictive in its assumptions.[6]

Finally, an interesting (and productive) branch of the literature examines the possibility of "linked negotiations" on transboundary pollution with other issues such as free trade

[5] Note that Barret (1999) analyzes the theory of international cooperation in the context of repeated games where players use contingent strategies, such as grim strategies.

[6] Schelling (2006) provides arguments about why there is no obvious formula to make the punishment fit the crime in IEAs.

agreements, which developed countries may use to achieve greater reductions in pollution by developing countries, see Whalley (1991) and Folmer *et al* (1993). Importantly, these models predict a limitation of the free-riding practices when the countries' interests are sufficiently complementary. In spite of their interest, these models have also been criticized because of: (1) the coercion they seem to recommend from developed nations to underdeveloped ones in order to induce better environmental practices among the latter, and (2) because of the difficulty to implement such limitations on real free-trade agreements, given the last advances of the World Trade Organization.

In order to overcome some of the shortcomings of the existing literature, in this paper I propose a model that limits countries' free-riding practices in certain cases (while it allows them under some parameter values) without the need to repeat the environmental game during different periods and without relying on social preferences among the countries. In addition, I do not need to allow the possibility of legal sanctions (or coercion in terms of trade agreements) to be enforced by the countries or by any international organization. These elements permit an easier analysis and complementary interpretations to the ones in the above models.

2.2.2 Literature on treaty compliance

There is substantial debate in the literature about how countries achieve compliance in international agreements. Chayes and Chayes (1995) argue that countries spend a lot of energy and time in preparing, drafting, negotiating and monitoring treaty obligations, which leads them to usually comply with their part of the treaty. Even though they recognize that noncompliance exists, they justify it by ambiguity of the treaty, the capacity limitation of status and uncontrollable social or economic changes. Moreover, they consider that sanctions are not necessary to ensure compliance. On the other hand, Downs et al. (1996) defend the idea that sanctions are an important element on treaty compliance. They argue that the evidence suggests that high levels of compliance and infrequent use of enforcement result from the low requirements of the agreement. Barret (1999) attempts to disentangle the debate. He concludes that the main constraint on international cooperation is free-riding deterrence, not compliance enforcement. This paper recognizes the fact that noncompliance plays an important role in the IEA, however the extreme case of complete violation of a treaty obligation is not observed in the equilibrium. Additionally, the absence of the full cooperative outcome in the results can mainly be explained by free-riding as in Barret (1999).

2.3 MODEL WHEN COUNTRIES CARE ABOUT THEIR OWN FULFILLMENT

Consider a two-stage complete information game. In the first stage of the game, the negotiation stage, countries decide the terms of the environmental agreement. In the second stage, the investment game, countries privately decide how much investment to make on emission-reducing technologies. Each country is endowed with w monetary units (e.g. governmental budget). Let x_i denote country i's monetary investments in clean technologies (alternatively in reduction of pollutant emissions), and let z_i represent its consumption of private goods. These private goods can be interpreted, generally, as the tax revenue raised by the government, which can now be kept for alternative uses in other expenditure programs not related with the IEA. Additionally, the marginal utility country i derives from alternative expenditure programs (private good) is one.

In particular, I assume that the difference between country i's actual investment in cleaner technologies, x_i, and the commitment level of investment that country i specified in the treaty, $c_i \geq 0$[7] (which is endogenously determined in the first stage of the game, section 3.3), represents the relative fulfillment of country i's commitment in the agreement. This difference can also be understood as the noncompliance cost. That is,

$$f_i = \alpha_i \left(x_i - c_i\right) \tag{2.1}$$

First, note that country i improves its opinion among green voters if its investment in cleaner technologies, x_i, is higher or equal than its commitment level, c_i; otherwise, if country

[7] Zero represents the case in which country i does not sign the agreement.

i invests less than what it was supposed to, $x_i < c_i$, green voters of country i perceive a lack of commitment to the agreement, which could lead them to penalize the incumbent party in future elections.[8] In addition, this difference is scaled by $\alpha_i \in [0, +\infty[$. In short, α_i[9] indicates the importance of green voters in country i. For instance, α_i can be interpreted as the percentage of politicians from green parties in the Senate or the percentage of population who belongs to environmental organizations. The higher is this percentage, the more negative is the impact of a lack of fulfillment of the agreement in the governments' utility. Finally, in further sections I consider an extension of this model, where countries are assumed to be concerned about each others' relative fulfillment of the IEA (instead of its own fulfillment).

Specifically, country i's utility function is represented by a quasilinear utility function, where private goods (money) enter linearly, while both total investments in clean technologies by country i and j, $G = x_i + x_j$, and country i's relative fulfillment of the terms in the IEA, f_i, are included in the nonlinear function $v(\cdot)$.

$$U_i (z_i, G, f_i) = z_i + v(G, f_i) \tag{2.2}$$

For simplicity, assume the nonlinear (concave) function $v(G, f_i) = \ln [mG + f_i]$[10]. Therefore, the representative country's maximization problem is given by

$$\max_{x_i} \; U_i (z_i, G, f_i) = z_i + \ln [mG + \alpha_i (x_i - c_i)] \tag{2.3}$$

$$\begin{aligned} \text{subject to } z_i + x_i &= w \\ x_i + x_j &= G \\ x_i &\geqslant 0 \end{aligned}$$

[8] This study assumes that green voters care about the total investment in clean technologies, G, and the compliance of their countries' agreement, f_i, but not about the commitment level signed in the IEA, c_i, since that level is not relevant per se in terms of achieving the objectives of the IEA.

[9] Note that this parameter can have different interpretations. For instance, in this model, it is understood as the countries concern about their green voters' penalization. However, it can also be interpreted as countries individualistic concerns about their self-image or moral motivations. Nyborg (2000) and Brekke et al. (2003) have developed interesting studies about moral motivation in public good games.

[10] The $ln(\cdot)$ function in country i's utility must be strictly positive, i.e. $mG + \alpha_i (x_i - c_i) > 0$.

Using $z_i = w - x_i$, and for a given value of $x_j \geqslant 0$ we can simplify the above program to

$$\max_{x_i} w - x_i + \ln\left[m(x_i + x_j) + \alpha_i (x_i - c_i)\right] \tag{2.4}$$

$$x_i \geqslant 0$$

In particular, the first term, $w - x_i$, represents the utility derived from the consumption of the remaining monetary units that have not been invested in clean technologies, i.e., that have not been invested in the public good. In the second term, m represents the return from the environmental good and $m(x_i + x_j)$ denotes the total return that country i obtains from the consumption of a higher level of environmental quality given its own investments, x_i, and the ones of country j, x_j. Finally, $\alpha_i (x_i - c_i)$ represents the return that country i derives from relatively fulfilling its commitment c_i in the environmental agreement or the cost that it incurs from the noncompliance of its agreements. Intuitively, an increase in country i's investment, x_i, has a *traditional public good* dimension, via $m(x_i + x_j)$, and an additional *fulfillment* dimension, via $\alpha_i (x_i - c_i)$.

2.3.1 Best response function

In order to gain a clearer intuition of the results, this subsection introduces country i's best response function, and the next section analyzes the optimal investment level. Henceforth, all proofs are included in the appendix.

Lemma 1

In the investment game with concerns about green voters, country i's best response function, $x_i(x_j)$, is

$$x_i(x_j) = \begin{cases} 1 + \frac{1}{m+\alpha_i}\left[\alpha_i c_i - m x_j\right] \text{ if } x_j \in [0, \frac{\alpha_i(1+c_i)+m}{m}[\\ 0 \text{ if } x_j \geqslant \frac{\alpha_i(1+c_i)+m}{m} \end{cases}$$

Let us compare the best response function $x_i(x_j)$ of a country which assigns a positive importance to the noncompliance cost, $\alpha_i > 0$, to that of a country which is *not* concerned about it, $x_i^{NC}(x_j)$, as in standard environmental models. In particular, when country i

assigns no importance to the population who cares about the relative fulfillment of IEA, $\alpha_i = 0$, country i's best response function becomes

$$x_i^{NC}(x_j) = \begin{cases} 1 - x_j \text{ if } x_j \in [0, 1[\\ 0 \ \text{ if } \ x_j \geqslant 1 \end{cases} \tag{2.5}$$

This expression is represented in figure 1, which helps in the comparison of the reaction functions. Specifically, note that $x_i(x_j)$ is always above $x_i^{NC}(x_j)$ for any x_j.[11] In other words, country i will always have higher levels of investment in emission-reducing technologies when it is concerned about green voters' punishment than otherwise, for any investments of country j, x_j.

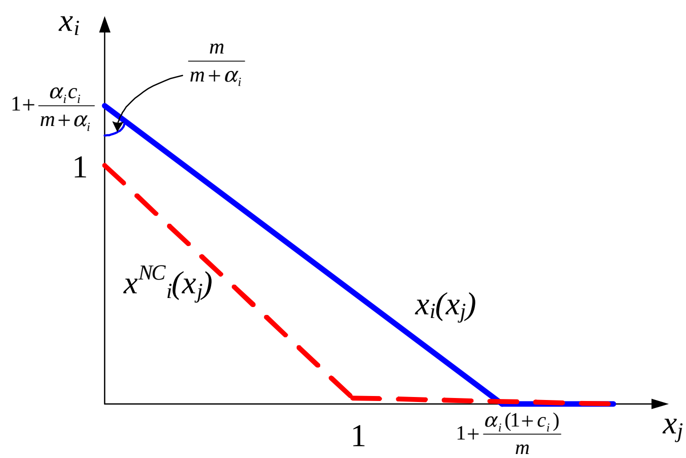

Figure 1: Model 1. Comparison Between BRFs

2.3.2 Non-cooperative equilibrium investments

In this section I examine the equilibrium strategies in the simultaneous Nash equilibrium resulting from both countries i and j applying lemma 1. The following proposition states the main result, and below I elaborate on its intuition and comparative statics.

[11]Note that $1 + \frac{\alpha_i(1+c_i)}{m} > 1$ for any parameter values, and any $c_i \geq 0$.

Proposition 1

In the investment game, every country i's Nash equilibrium investment in emission-reducing technologies is

$$x_i^* = \begin{cases} 1 + \frac{\alpha_i c_i}{m+\alpha_i} \textit{ if } \alpha_i > \bar{\alpha}_i(\alpha_j) \\ \frac{\alpha_i(1+c_i)(\alpha_j+m)-\alpha_j m c_j}{\alpha_j m+\alpha_i(\alpha_j+m)} \textit{ if } \alpha_i \in]\hat{\alpha}_i(\alpha_j), \bar{\alpha}_i(\alpha_j)] \\ 0 \textit{ if } \alpha_i \in]0, \hat{\alpha}_i(\alpha_j)] \end{cases}$$

where $\bar{\alpha}_i(\alpha_j) = \frac{\alpha_j c_j m}{(1+c_i)(\alpha_j+m)}$ *and* $\hat{\alpha}_i(\alpha_j) = \frac{mc_j+\alpha_j(1+c_j)(m+c_j)}{(1+c_i)m}$

In particular, country i's investment in clean technologies is at its maximum level when its concern about green voters, α_i, is sufficiently high, i.e., $\alpha_i > \bar{\alpha}_i(\alpha_j)$. When the importance that country i assigns to green voters, α_i, decreases below $\bar{\alpha}_i(\alpha_j)$, its optimal investment also decreases, as the above proposition shows. That is, country i is not highly concerned about its own relative fulfillment of the IEA because it does not perceive the group of green voters as being relevant in future elections. Finally, if α_i drops below the threshold $\hat{\alpha}_i(\alpha_j)$, then its concerns about green voters' punishment are not strong enough to support any positive investment in clean technologies from country i. Hence, from proposition 1 we can conclude that the full free-riding outcome is not a solution of this game when the weight that the country assigns to green voters is above a particular threshold, as the following corollary specifies.

Corollary 1

In the investment game, every country i's Nash equilibrium investment in emission-reducing technologies, x_i^**, is strictly positive, if and only if* $\alpha_i > \hat{\alpha}_i(\alpha_j)$.

Additionally, the following lemma indicates under what parameter values we can expect countries' *aggregate* investment in clean technologies to be higher than their investment when countries are not concerned about their own relative fulfillment of the IEA.

Lemma 2

In the investment game, when one of the countries is sufficiently concerned about green voters' punishment, i.e., $\alpha_i > \hat{\alpha}_i(\alpha_j)$ *or* $\alpha_j > \hat{\alpha}_j(\alpha_i)$, *the aggregate Nash equilibrium investment in emission-reducing technologies , G, is greater than one for any parameter values.*

Thus, as long as at least one of the countries is sufficiently concerned about the noncompliance cost, i.e., $\alpha_i > \hat{\alpha}_i(\alpha_j)$ or $\alpha_j > \hat{\alpha}_j(\alpha_i)$, the total optimal investment in the simultaneous environmental game, $G = x_i^* + x_j^*$, is higher than the total investment obtained in standard environmental games in which countries do not assign any weight to the relative fulfillment of the IEA. Let us now examine how country i's optimal investment in clean technologies changes in the different parameters of the model. The following lemma summarizes these comparative statics about x_i^*, while the discussion below elaborates on its intuition.

Lemma 3

In the environmental game of investment in emission-reducing technologies, country i*'s equilibrium investment level,* x_i^*, *is weakly increasing (decreasing) in* c_i (c_j), *for any parameter values.*

Hence, country i's equilibrium investment, x_i^*, is increasing in the (non-enforceable) own commitment, c_i, that country i accepts when signing the IEA. The increase in x_i^* is due to country i's own obligation to fulfill the contract, given that its lack of fulfillment can be punished by voters with strong environmental concerns. Interestingly, an increase in country j's commitment of investment in the IEA, i.e., an increase in c_j, reduces country i's optimal investment in clean technologies, x_i^*. This result can be explained by the fact that a higher commitment of country j in the IEA "relaxes" country i's incentives to invest in clean technologies. Let us now examine how x_i^* varies in country i and j's concerns about green voters.

Lemma 4

In the environmental game of investment in emission-reducing technologies, country i*'s equilibrium investment level,* x_i^*, *is weakly increasing (decreasing) in* α_i (α_j).

First, note that country i's Nash equilibrium level of investment, x_i^*, is increasing in its own concern about green voters, α_i. Thus, if green voters represent an important proportion of the population who can affect the future elections results, then country i will invest higher levels of x_i in order to fulfill its commitment, c_i. On the other hand, country i decreases its investment, x_i, if the importance of green voters in other countries, α_j, increases. Clearly, country i knows that an increase in α_j induces country j to achieve a greater fulfillment of its own commitments, increasing x_j, what ultimately reduces country i's investment since country i's best response function is negatively sloped.

2.3.3 Equilibrium proposals

The previous section analyzed the optimal investment levels of each country, *given* a specific commitment of investing in clean technologies specified in the IEA for every country, c_i and c_j. This section goes one step back and, using sequential rationality, examines the optimal investment levels that every country accepts for itself and the other country –the pair (c_i, c_j)– in the subgame perfect Nash equilibrium (SPNE) of the game describing the negotiation and posterior implementation of the IEA. The following proposition specifies countries' incentives in this negotiation stage of the game, and below I discuss the SPNE strategies of the signatory countries in the IEAs.

Proposition 2

Every country i's equilibrium payoff from playing the investment game, with the optimal investments $\left(x_i^*, x_j^*\right)$ *determined in proposition 1, is weakly decreasing (increasing) in* c_i (c_j).

Hence, in the negotiation stage of the game –where countries determine the investment levels (c_i, c_j) to be included in the text of the IEA– every country i has incentives to set low environmental standards for itself (low c_i), but high requirements for other countries (high c_j). This result has important consequences in the optimal proposals of the pair (c_i, c_j) to be voted during the negotiation stage under any voting procedure, since countries want to shift the greatest burden of the IEA to other countries.

The determination of the commitment levels that each country signs in the agreement could be obtained through the Nash bargaining solution concept. However, it cannot be applied in the model because the model's payoff structure is not well-behaved.

Therefore, the negotiation stage is represented by a one-shot game in which country i proposes a pair of commitment levels (c_i, c_j) and country j is allowed to either accept or reject such proposal. This game is known in the game theory literature as the ultimatum bargaining game. In particular, the following proposition analyzes the SPNE strategies resulting from a fairly simple voting procedure.

Proposition 3

If the voting procedure is represented by the ultimatum bargaining game, the equilibrium investment level for every country i is,

First case: if $\alpha_i > \bar{\alpha}_i(\alpha_j)$ and $\alpha_j \in]0, \hat{\alpha}_j(\alpha_i)]$

$(c_i^*, c_j^*) = (\frac{\alpha_i(-1+m+\alpha_i)}{1+\alpha_i}, \frac{e^w(\alpha_i^3 e^w m + mB + \alpha_i(1+m)B + \alpha_i^2(-1+e^w m^2)}{\alpha_j(1+\alpha_i)(m+\alpha_i)})$

$(x_i^*, x_j^*) = (\frac{m+\alpha_i(1+m+\alpha_i(m+\alpha_i))}{(1+\alpha_i)(m+\alpha_i)}, 0)$

Second case: if $\alpha_i \in]\hat{\alpha}_i(\alpha_j), \bar{\alpha}_i(\alpha_j)]$ and $\alpha_j \in]\hat{\alpha}_j(\alpha_i), \bar{\alpha}_j(\alpha_i)]$

$c_i^* \in [0, \frac{\alpha_i(-1+w)+m(-1+2w)+\Psi \log \Psi + m \log \Gamma)}{\alpha_i}]$ and

$c_j^* \in [\frac{-\alpha_i\alpha_j + \alpha_j m(-1+w) + \alpha_i \Gamma w + (\alpha_j m + \alpha_i \Gamma) \log \Gamma}{\alpha_j \Psi}, \frac{\alpha_j(-1+w)+m(-1+2w)+\Gamma \log \Gamma + m \log \Psi)}{\alpha_j}]$

$x_i^* \in [\frac{\Psi - mw - m \log \Gamma}{\Psi}, w + \log \Psi]$ and $x_j^* \in [\frac{\alpha_j(2\Psi - mw - m \log \Gamma)}{\alpha_j m + \alpha_i \Gamma}, w + \log \Gamma]$

Third case: if $\alpha_i \in]0, \hat{\alpha}_i(\alpha_j)]$ and $\alpha_j > \bar{\alpha}_j(\alpha_i)$

$c_i^* \in [0, \frac{e^{-w}(-1+e^w mw + e^w m \log[\alpha_j + m])}{\alpha_i}]$ and $c_j^* = \frac{(m+\alpha_j)(-1+w+\log[\alpha_j+m])}{\alpha_j}$

$x_i^* = 0$ and $x_j^* \in [1, w + \log[\alpha_j + m])$

Notice that country i exerts all its proposing power, since country j accepts any proposal leading to zero payoffs. Moreover, the negotiation process can be represented by three different cases which depend on the concern levels of the proposing country and the country receiving the offer. Therefore, the agreed commitment levels vary in each case, for instance, they depend in what the concern level of the first mover is (which has all the bargaining power) and how important the noncompliance cost is for the country which accept or reject such offer. Additionally, in the event of a rejection from country j, the total commitment levels are equal to zero and both countries get a zero payoff. In other words, if the environmental agreement is not signed countries will not benefit from the environmental quality than they could have achieved otherwise.

Of course, in more elaborated settings, such as the voting procedures developed in different international organizations, country i cannot take full advantage of its proposing power. Instead, it may propose less extreme investment pairs (c_i, c_j), since the possibility of playing some cooperative-punishment strategy in this repeated game might induce higher payoffs for country i. In spite of other considerations, the model that this paper analyzes can clearly capture countries' incentives during the negotiation of the IEA, which emphasizes countries efforts in shifting the greatest possible burden of the (non-binding) commitments included in the IEA to other countries. Moreover, proposition 3 suggests that even in the extreme case where country i's commitment level is zero (it can be interpreted as not signing the treaty), it will still invest in clean technologies as long as country i is concerned enough about green voters.[12]

Corollary 2

In the first case of the negotiation stage of the environmental game where country i proposes a pair $(0, \frac{e^w(-1+e^w m)}{\alpha_j})$ and country i's concern level is $\alpha_i \leq 1 - m$, country i's optimal investment in clean technologies is strictly positive.

Hence, the extreme situation where country i exerts all its bargaining power in the nego-

[12] Note the connection of these results with those in the literature on strategic pre-commitment, as in Fudenberg and Tirole (1984) and Balboa et al. (2004). Indeed, in this literature players choose the level of an irreversible variable, such as physical capital or tax, during the first stage with the objective to influence the strategic environment of the game played during the second stage. Similarly, in this model every country i uses the negotiation stage of the IEA to reduce its own commitment level (since this reduces its non-compliance costs in the second stage), and increases the other country's commitment level (given that this leads the other country to invest more in clean technologies during the second stage). In summary, every country uses the negotiation stage in order to shift most of the burden of the public good provision to the other country in the second stage.

tiation stage (zero commitment level) is compensated in the investment game. Since, when country i's concern level is strictly positive it invests positive amounts in clean technologies. Finally, notice that when the utility from the total investment in clean technology, m, increases, then country i proposes a less extreme pair of commitment levels, i.e. $c_i^* > 0$.

Corollary 3

In the negotiation stage of the environmental game where country i proposes a pair (c_i^*, c_j^*) *and country i is very (relatively) concerned about the noncompliance cost and country j is not (relatively) concerned,* c_j *is increasing in* α_i.

The above corollary represents country i's interest in imposing high commitment levels on country j. If the proposer suffers a high political cost when it does not fulfill its environmental agreements, then it has incentives in offering a positive c_j (even though the responder does not comply its agreement in the second stage of the game).

Corollary 4

In the negotiation stage of the environmental game where country i proposes a pair (c_i, c_j) *and country i and j are relatively concerned about the noncompliance cost ,* G^* *is increasing in* α_i and α_j.

In other words, when signatory countries are concerned about the noncompliance cost, the total investment in clean technology obtained in the investment game increases. That is, if countries which participate in the negotiation of the IEA have high political costs, in terms of noncompliance of their environmental agreements, it positively affect the results of the treaty.

Finally, figure 14 in the appendix depicts the relationship between country i and j's concern levels. Specifically, region 2 shows that every country tries to impose the most demanding environmental standard on the other country but not on themselves during the negotiation stage. Notice that country i could propose zero commitment level for itself (avoiding any political cost), however this is not the case when its concern level lies in region 2. Specifically, country i and j sign positive commitment levels which can be explained by the traditional public good dimension. In fact this result reflects the trade off between countries' political cost and the total return from the total investment in clean technology.

2.4 A MODEL WHEN COUNTRIES CARE ABOUT EACH OTHERS' FULFILLMENT

Let us now consider the model presented in section 3, but assuming that country i's concerns about the relative fulfillment of the IEA depends on the extent to which country j fulfills its commitment in the agreement, $c_j \geqslant 0$. That is, country i's maximization problem is now defined as

$$\max_{x_i} \quad w - x_i + \ln\left[m(x_i + x_j) + \alpha_i\left(x_j - c_j\right)\right] \tag{2.6}$$

where country i improves his perception of country j's serious commitment of carrying out the treaty if country j's investment in cleaner technologies, x_j, is higher than or equal to his commitment level, c_j; otherwise, if country j invests less than what it was supposed to, $x_j < c_j$, country i perceives a lack of commitment in the fulfillment of the agreement, which leads to a negative perception from country j's actions. Additionally, α_i indicates the importance that country i assigns to country j's fulfillment of its part of the agreement, where as before $\alpha_i \in [0, +\infty[$.

Intuitively, note that in this model an increase in country j's investment, x_j, imposes both a positive direct and indirect effect on country i's utility level. The positive *direct* effect from x_j on country i's utility is just the usual one arising from the public good nature of country j's investment on emission-reducing technologies. Country j's investments, additionally, impose a positive *indirect* effect on country i since these investments increase the relative commitment that country i perceives from country j's actions, i.e., higher x_j increases $\alpha_i\left(x_j - c_j\right)$, for any given commitment level c_j. The following lemma describes country i's best response function in this context.

Lemma 5

In the simultaneous environmental game of investment in emission-reducing technologies with concerns about the each others' fulfillment of the international agreement, country i's best response function, $x_i(x_j)$, *is*

$$x_i(x_j) = \begin{cases} \frac{1}{m}\left[\alpha_i c_j + m - (\alpha_i + m)x_j\right] \text{ if } x_j \in [0, \frac{\alpha_i c_j + m}{\alpha_i + m}[\\ 0 \text{ if } x_j \geqslant \frac{\alpha_i c_j + m}{\alpha_i + m} \end{cases}$$

A comparison between the best response function $x_i(x_j)$ of a country concerned about other country's fulfillment of the agreement, $\alpha_i > 0$, with respect to that of a country which is *not* concerned about the other country's commitment with the treaty, $x_i^{NC}(x_j)$ will give us more intuition about the countries' strategic behavior.

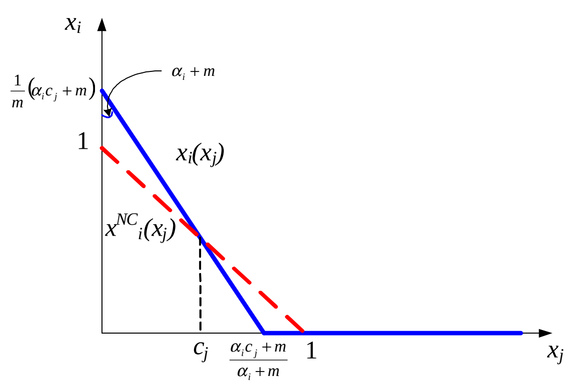

Figure 2: Model 2. Comparison betwee BRFs

Figure 2 shows that $x_i(x_j)$ is steeper than $x_i^{NC}(x_j)$ for any x_j. In particular, $x_i(x_j)$ is above $x_i^{NC}(x_j)$ for any $x_j < c_j$, and below otherwise. In other words, country i compensates country j's investments when it is below its country's commitment in the IEA, c_j. Specifically, note that when country j's actual investment in emission-reducing technologies is lower than the level it signed in the IEA, $x_j < c_j$, country i experiences a disutility from the lack of commitment it interprets from country j's actions. Hence, in order to compensate for such low level of investment country i invests more than it would in the case of not being concerned about the fulfillment of the contract. In contrast, when $x_j > c_j$, country i experiences an increase in its utility level given the strong commitment with the IEA from country j. In this case, country i makes an optimal investment in clean technologies below that it would carry out when not being concerned about the performance of the contract.

Let us now examine the equilibrium strategies in the simultaneous Nash equilibrium resulting from both countries i and j applying the above best response function. The following proposition states the main result, and below I elaborate on its intuition.

Proposition 4

In the negotiation game, every country i's Nash equilibrium investment in emission-reducing technologies is

$$x_i^* = \begin{cases} \frac{\alpha_i c_j + m}{m,} \ \textit{if} \ \alpha_i > \bar{\alpha}_i(\alpha_j) \\ \frac{\alpha_j c_i m + \alpha_i(\alpha_j c_i + m - c_j m)}{\alpha_j m + \alpha_i(\alpha_j + m)} \ \textit{if} \ \alpha_i \in]\hat{\alpha}_i(\alpha_j), \bar{\alpha}_i(\alpha_j)] \\ 0 \ \textit{if} \ \alpha_i \in]0, \hat{\alpha}_i(\alpha_j)] \end{cases}$$

where $\bar{\alpha}_i(\alpha_j) = \frac{\alpha_j(c_i - m) + m - m^2}{c_j(\alpha_j + m)}$ *and* $\hat{\alpha}_i(\alpha_j) = \frac{\alpha_j c_i m}{m(c_j - 1) - \alpha_j c_i}$

In particular, country i's investment in clean technologies is at its maximum level when country i's concerns about other countries fulfillment of the agreement, α_i, is sufficiently high, i.e., when $\alpha_i > \bar{\alpha}_i(\alpha_j)$. However, if the importance that country i assigns to country j's commitment with the contract, α_i, decreases below $\bar{\alpha}_i(\alpha_j)$, its optimal investment also decreases, as the above proposition shows. Finally, when α_i drops below the threshold $\hat{\alpha}_i(\alpha_j)$, then its concerns about country j honoring the IEA are not strong enough to support any positive investment in clean technologies from country i. Let us now analyze under what parameter values the aggregate investments in clean technologies are above those in standard environmental games.

Lemma 6

In the negotiation game, the aggregate Nash equilibrium investment in emission-reducing technologies, G, is greater than one if and only if either $\alpha_i > \bar{\alpha}_i(\alpha_j)$ *or* $\alpha_j > \bar{\alpha}_j(\alpha_i)$*. Additionally, when neither* $\alpha_i > \bar{\alpha}_i(\alpha_j)$ *nor* $\alpha_j > \bar{\alpha}_j(\alpha_i)$ *are satisfied, G is greater than one if and only if* $c_i + c_j > 1$*.*

As in the previous model, the total optimal investment in the simultaneous environmental game, $G = x_i^* + x_j^*$, is greater than one when at least one of the countries is highly concerned about each others' relative fulfillment, i.e., $\alpha_i > \bar{\alpha}_i(\alpha_j)$ or $\alpha_j > \bar{\alpha}_j(\alpha_i)$. However, when both countries are not highly concerned about each others' fulfillment, total investments are only higher than those in standard environmental games when the total environmental goals specified in the IEA are relatively demanding, i.e., when $c_i + c_j > 1$. Regarding the comparative statics of the Nash equilibrium investment levels, x_i^*, the following lemma confirms that we can extend our intuitions from the previous section.

Lemma 7

In the negotiation stage in which countries are concerned about each others' relative fulfillment of the IEA, both corollary 1 and lemma 3 hold.

Hence, in this setting lemma 3 can be interpreted as follows. Country i's equilibrium investment, x_i^*, is increasing in the (non-enforceable) own commitment, c_i, that it accepts when signing the IEA. Interestingly, the increase in x_i^* is not due to country i's own obligation to fulfill the contract (as in previous sections), but instead, because a higher commitment of country i in the IEA "relaxes" country j's incentives to invest in clean technologies. Indeed, since now country i is supposed to invest more (higher c_i), country j invests less (lower x_j^*), which ultimately leads country i to increase its investment to compensate country j's lack of investment in emission-reducing technologies. Similarly, an increase in country j's agreed level of investment in the IEA –i.e., an increase in c_j, as in the second result of the above lemma– reduces country i's optimal investment in clean technologies, x_i^*. In the following lemma I examine how x_i^* varies in country i and j's concerns about each others' fulfillment of the environmental agreement.

Lemma 8

In the environmental game of investment in emission-reducing technologies, country i's equilibrium investment level, x_i^*, is weakly decreasing (increasing) in α_i (α_j), if and only if $c_j > 1 - c_i$.

First, note that country i's Nash equilibrium level of investment, x_i^*, is *decreasing* in its own concern about country j's fulfillment of the contract's requirements, α_i, if and only if $c_j > 1 - c_i$. This result is opposed to that we obtained in lemma 4. In particular, it specifies that if the commitment of investment in emission-reducing technologies that country j signs in the IEA is sufficiently high, then country i perceives its investment as less necessary, similarly to the above discussion about the effects of an increase in c_j. Otherwise, if $c_j < 1 - c_i$, then country i increases its investment as α_i increases, since it considers that country i must compensate country j's lack of investment in clean technologies.

An alternative interpretation of this result would focus on how "demanding" are the environmental goals included in the IEA. When the investment objectives specified in the IEA are extremely demanding, i.e., $c_i + c_j > 1$, then country i's optimal investment in emission-reducing technologies decreases in their own concern, α_i, about country j's fulfillment of the contract. In contrast, international agreements with conservative goals, $c_i + c_j < 1$, make country i's investment in clean technologies to be increasing in α_i.[13]

Let us finally analyze what happens with country i's optimal investment, x_i^*, when the importance that country j assigns to country i's fulfillment of the contract requirements, α_j, increases. In particular, an increase in α_j raises country i's equilibrium investment, x_i^*, if and only if $c_j > 1 - c_i$. Clearly, now if country j assigns a greater importance to country i's fulfillment of the contract and the investment level that country i specified in the IEA is

[13] This interpretation is related to the results obtained by Barrett (1994a) and Downs et. al. (1996). When the agreement establishes low requirements (or gains to cooperate are small) free riding behavior is less preeminent. Therefore, a highly concerned country will exert higher efforts to achieve the compliance of the agreement.

relatively high (i.e., $c_j > 1 - c_i$ is equivalent to $1 - c_j < c_i$) leads country j to reduce x_j, increasing x_i^* as a consequence.

Similarly to the previous intuition, if we interpret these results in terms of how demanding are the goals of the IEA, one can conclude that when the IEA is very demanding (conservative) country i's investments in emission-reducing technologies are increasing (decreasing) in the importance that other countries assign to country i's fulfillment of the contract, α_j. Finally, we can briefly analyze the negotiation stage of the IEA given the above optimal investment levels for every country.

Proposition 5

Every country i's equilibrium payoff from playing the investment game, with the optimal investments $\left(x_i^*, x_j^*\right)$ *determined in proposition 3, is weakly decreasing (increasing) in* c_i *(*c_j*).*

The above proposition is indeed equivalent to proposition 2, and in this context it emphasizes countries' incentives to shift most of the burden of the IEA to other countries, trying to make certain that the IEA specifies high commitment levels for other countries, c_j, and low for themselves, c_i. Finally, the following proposition defines the subgame perfect Nash equilibrium of the game under the assumption that countries are concerned about each others' relative fulfillment of the IEA.

Proposition 6

If the voting procedure is represented by the ultimatum bargaining game, the equilibrium investment level for every country i is,

First case: if $\alpha_i > \bar{\alpha}_i(\alpha_j)$ and $\alpha_j \in]0, \hat{\alpha}_j(\alpha_i)]$

$(c_i^*, c_j^*) = (\frac{e^{-w}(m - \alpha_i m + e^w \Gamma(\alpha_i^2 + (-1 + \alpha_i + \alpha_i^2) m))}{(-1+\alpha_i)\alpha_j m}, \frac{\alpha_i(1+m)}{(-1+\alpha_i)})$

$(x_i^*, x_j^*) = (0, \frac{\alpha_j m c_j + m\Gamma + \sqrt{4\alpha_j^2 \Gamma^2 + m^2(\Gamma - \alpha_j c_j)^2}}{2m\alpha_j \Gamma})$

Second case: if $\alpha_i \in]\hat{\alpha}_i(\alpha_j), \bar{\alpha}_i(\alpha_j)]$ and $\alpha_j \in]\hat{\alpha}_j(\alpha_i), \bar{\alpha}_j(\alpha_i)]$

$c_i^* \in [0, \frac{\alpha_j w + m(-1+2w) + \lambda \log m)}{\alpha_j}]$, $c_j^* \in [\frac{-\alpha_j m + \alpha_i m w + \alpha_j \Psi w + (\alpha_j m + \alpha_i \Gamma) \log m)}{\alpha_i \Gamma}, \frac{\alpha_i w + m(-1+2w) + \varphi \log m)}{\alpha_i}]$

$x_i^* \in [\frac{m - mw - m \log m}{\Gamma}, w + \log m]$ and $x_j^* \in [\frac{-m^2 + \Gamma\varphi w + \Gamma\varphi \log m}{\alpha_j m + \alpha_i \Gamma}, w + \log m]$

2.4.1 Discussion and Applications

2.4.2 Discussion on countries' asymmetric fulfillment of IEAs

Both of the models presented in this paper clearly predict that countries invest (weakly) higher levels in emission-reducing technologies than in standard environmental games. Interestingly, in the first model, the increase in country's investment is due to its concern about its own relative fulfillment of the IEA, whereas in the second model this increase is due to the country's concern about other countries' relatively fulfillment of the agreement. Notwithstanding their differences, their similar predictions can explain why certain countries fulfill to a great extent the commitments they acquire when signing IEAs, even if these commitments are relatively informal and not perfectly enforceable.

Additionally, both models predict that a country's optimal investment decision, x_i, increases in the country's own commitment level, c_i, and decreases in other countries' commitments, c_j. This result indicates that both countries "relax" their optimal investments when other countries' commitments in the IEA increase, this can rationalize why different countries condition their investment decisions on other cosigners' particular commitments in the IEA, despite knowing that such commitments are non-binding and may not be implemented by the cosigners of the treaty.

Finally, note that the main difference between the results of both models is on the comparative statics of countries' equilibrium investment, x_i^*, with respect to α_i and α_j. In the case that countries are concerned about their own relative fulfillment of the agreement, their equilibrium investments are clearly increasing in the importance that they assign to their own fulfillment of the IEA, α_i, and decreasing in the weight that the other country assigns, α_j. In contrast, when countries are concerned about each others' relative fulfillment of the environmental agreement, their equilibrium investments move in opposite directions: decreasing in the importance every country assigns to other countries fulfillment of the IEA, α_i , and increasing in the weight that other countries assign, α_j, if and only if the agreement is extremely demanding.

The last result may explain the perspective on compliance of Downs et al. (1996), (see section 2.2). They argue that when countries sign low commitments levels in the IEA, it is very likely to observe that signatories fulfill their compromises. Hence, in this model, low agreed levels will induce countries which are concerned about other's countries fulfillment to comply its commitment levels. The findings rationalize why countries prefer to adopt agreements that state feasible and realistic commitment levels.

2.4.3 Applications to international environmental negotiations

Regarding the negotiation stage, I first show that countries' equilibrium payoff from playing the investment game is weakly increasing in c_j and decreasing in c_i, for both models developed in this paper. Hence, regardless of the voting procedure which finally decides which levels of c_i and c_j are included in the IEA, countries clearly prefer to shift most of the commitments of investment in clean technologies to other countries. Additionally, in the particular case in which the voting procedure is similar to the ultimatum bargaining game, I show that country i uses its proposer power to reduce c_i (and increase c_j) as much as possible. Ex ante, this could make us conclude that countries are leading their negotiations towards a situation in which they all want to free-ride each other's efforts in emission-reducing technologies, without bearing any costs. However, this is not the case, as previous sections show. Specifically, every country's optimal investment in clean technologies is strictly higher than zero (both when countries are concerned enough about green voters' punishment and when they are concerned about each others' fulfillment), and increases in certain parameters. In particular, this is true for the country which proposed the investment pair (c_i, c_j) specified in the international agreement, as well as for the country which accepted such proposal.

Hence, the behavior initially predicted for the voting stage –which one could describe as free-riding of the country with the greatest bargaining power– is then compensated by the second stage of the game, where countries decide how much to invest in emission-reducing technologies, since no country decides to operate as a pure free-rider given their mutual concerns about the fulfilling of the IEA, $\alpha_i, \alpha_j > 0$, as opposed to the prediction of the model when countries do not assign any importance to such fulfillment of the agreement,

$\alpha_i = \alpha_j = 0$. The negotiation process can rationalize some cases that are observed in current IEAs. For instance, it reflects the EU interest in requiring the participation of India or China in the Kyoto protocol. In particular, the EU is willing to increase its commitment level (30% reduction in emissions) if and only if countries which are considered heavy pollutants sign the agreement. Finally, the results can explain the United States' case in the Kyoto protocol, where in terms of the model, U.S. commitment level is zero (U.S. did not ratify the protocol in the Senate). However, any positive investment in clean technologies would be interpreted by U.S. concerns about green voters.

These results can shed some light on some relatively surprising real-life cases of IEAs, where different countries first need long periods of time in order to reach an agreement about how much each country will reduce its emission of pollutants (or alternatively, how much resources to invest in clean technologies). In particular, countries usually want to impose important quotas on other countries (high values of c_j), but at the same time are reluctant to determine high quotas for themselves, c_i. If these international agreements were perfectly enforceable, countries would have a strong incentive to fight for a favorable division of environmental quotas. These agreements are, however, clearly not perfectly enforceable, what limits the possibility of rationalizing such lengthy negotiations from the perspective of perfectly enforceable quotas.

Indeed, this model predicts that countries fight for such favorable quotas not because they do not want to implement high investments in clean technology which is then benefited by other countries given its public good nature. Instead, this model predicts such negotiations because, even if the investment pair (c_i, c_j) included in the IEA is not enforceable, it enters as a reference point[14] in the countries' utility function. In particular, as previous sections show, country i's optimal investment decreases in c_j (and increases in c_i). However, since these investments are costly, countries want to specify the contents of the IEA such that it sets high environmental standards for other countries (high c_j) and low requirements for themselves (low c_i).

[14] A deeper analysis of the effect of reference points on players' strategic incentives in sequential move games can be found in Espinola-Arredondo and Munoz-Garcia (2007).

2.5 CONCLUSIONS

This paper analyzes a two-stage game where countries, first, decide the pair of investment levels (c_i, c_j) in emission-reducing technologies to be included in an international environmental agreement with no enforcing possibilities. Then, in the second stage, every country independently and non-cooperatively determines how much to invest in this technology, given its character of public good, and given the country's concern about the relative fulfillment of the international agreement, either by itself or by other countries.

The study shows that, first, every country's investment level in clean technology is nonzero for most parameter values, unlike those models analyzing environmental games in which players (countries) are not concerned about the relative fulfillment of the contract's requirements. In addition, it examines how optimal investments vary in different parameters. For instance, the country's equilibrium investment in clean technologies can actually increase in the importance (or political representation) of green voters. Similarly, this paper also shows that these investments increase in the country's concerns about other countries' relative fulfillment of the IEA if, in addition, this IEA sets relatively low emission reduction goals. In contrast, if the IEA sets high goals, it shows that a country's investment in clean technologies decrease in the country's concerns about other countries' relative fulfillment of the IEA. This result supports many real-life observations, in which countries prefer to specify low goals (instead of unrealistic levels) of environmental improvement to be included in the IEA. Finally, the study considers how countries' investment varies in the commitment that every country acquires in the IEA, discussing why this result does not necessarily depend on the agreement (since it is non-enforceable), but instead on the countries' own incentives in the environmental game.

A crucial element in this model is the negotiation stage of the game, where countries decide the investment levels to be included in the IEA. This paper analyzes the case when the voting procedure resembles that in an ultimatum bargaining game, that is, country's optimal promises prescribe that all the (non-enforceable) investment in clean technologies is carried out by other countries, leaving no investment burden for itself. The results suggest that, in spite of that these commitment levels are non-enforceable, some countries invest positive amounts in environmentally oriented technology, even those countries who suggest the radical proposal $(c_i^*, c_j^*) = (0, c_j)$ during the voting stage of the game. The findings predict that is more likely that an agreement will be forthcoming if the participating countries are relatively concerned about the noncompliance cost. Additionally, increases in the advertising of countries' fulfillment of their environmental agreements by organizations such as the United Nations would raise countries' concern level.

Finally, many results of this model permits a more general rationalization of real-life practices during (and after) negotiations of IEAs. First, they support lengthy discussions during the approval stage of IEA in which every country wants to get a favorable division of the proposed investments in clean technologies that the agreement specifies, even when the IEA is clearly non-enforceable. Second, they explain why different countries do fulfill the commitment they acquired when signing an IEA, while others do not; and how this strategy depends on certain parameter values, such as how demanding is the international agreement, or the international orientation of these countries' media services.

Different extensions would clearly enrich the analysis of this general model. First, the paper develops a two stage complete information game. However, it would be interesting to analyze the case in which each country has private information about its concern level, α_i. Hence, in this setting country i can send a message about how much it cares about green voters or other countries' fulfillment of the agreement through its commitment level in the IEA, c_i. Second, countries' utility (or disutility) only comes from own or other countries' fulfillment of the contracts' requirements, while they do not consider, for example, their own bad "international image" from not fulfilling the terms of the IEA, which also could be included in a more general model. Moreover, it would be interesting to study the case where voters are represented by a different utility than that of the government. Finally, I assumed that all countries obtained the same utility from the global environmental quality, however a relaxation of this assumption allows us to analyze the role of this variable in the negotiation game, enriching the previous analysis. Further research in this area would enhance and clarify our understanding of countries' incentives in the negotiation and (partial) implementation of international agreements involving global public goods.

3.0 GREEN AUCTIONS: A BIODIVERSITY STUDY OF MECHANISM DESIGN WITH EXTERNALITIES

3.1 INTRODUCTION

Many governments face the problem of how to procure the implementation of a project considering the externalities that it produces in its surroundings. For instance, when the government procures the construction of a bridge, it has to consider the cost of this project and the bridge's characteristics but also the externalities that it will produce in all the community that will use it. These considerations about externalities are especially present when the government wants to develop a conservation program that increases the level of biodiversity of some particular territory. In particular, this situation exists when the government does not develop the environmental project by itself, but assigns its implementation to the landholders.

The paper assumes that landholders dedicate their land to market activities before the project is implemented. As a consequence, the only tool that the government can use to induce landholders' participation in the conservation program is through a monetary compensation. Therefore, the government has to achieve a balance between the costs of implementing the project and the biodiversity that it generates. A relevant characteristic in this kind of situation is related with the fact that the landholder has more information than the government about his costs and the resources needed to implement a biodiversity project in his land. Hence, we are dealing with information asymmetry, which can provoke opportunistic behavior from the landholder. For example, he can hide information about his real costs or about the land use intensities. As a result, the government could be overcompensating this opportunistic landholder. In this respect, auctions reduce the scope for opportunistic behavior resulting from informational asymmetries between the seller and the bidders. Specifically, in this type of procurement auctions the government announces a contract for the assignment of a biodiversity project in a specified area and calls for bids from potential bidders (landholders).

The government is not only interested in improving the biodiversity, however, but it also has to optimize the monetary resources coming from all society, as noted above. This implies that the landholder submits a multidimensional message in this mechanism, which is composed of the cost of the project and a non-monetary variable that represents the biodiversity improvement that the project will generate. In this sense, one of the main objectives of this paper is to construct an optimal mechanism that maximizes the government's objective function and induces landholders to participate in the conservation program. Once this optimal mechanism is found, it can easily be transformed into a procurement auction by appropriately applying the optimal allocation and payment rule.

Some examples of procurement auctions used to protect the biodiversity of a specific region are found in the United States and Australia. In the case of the United States, the Department of Agriculture has implemented auctions in which the landholders bid for payments for retiring their lands from farm production for a period of 10 to 15 years. In Victoria (Australia) the government has developed an auction of conservation programs where the landholders bid for payments they can receive for undertaking conservation activities.

Additionally, the conservation program described above also needs to consider the existence of externalities[1] derived from the project. I assume that the landholder producing biodiversity generates externalities for those neighboring landholders who decide not to participate. An important assumption in this study is that the environmental improvement –understood as an increase in the biodiversity level in the land– causes a positive (negative) externality in some specific area. Furthermore, this externality leads to a reduction (increase)

[1] An externality exists whenever the welfare of some agent, landholder in this model, depends on activities under the control of some other agent.

in the cost of all the landholders who do not produce biodiversity and keep using their land in the same activities as before the conservation program. Since these activities were supposed to have only market concerns I refer to these activities carried out by landholders before the procurement auction as "market activities".

For simplicity, I design a mechanism that introduces the externalities generated by an environmental project into the cost function of every nonparticipant landholder. The first part of the paper is focused on determining the consequences of having only a *negative* externality in the design of the mechanism. An example of negative externality can be found in the procurement of a project to increase the species of oaks. Let us consider two landholders, one of them producing honey and the other one dedicating his land to the production of flowers, and both of them sell their products in the market. Given this initial situation, the government calls for the procurement of a project to increase the population of oaks. As a result, the producer of flowers dedicates his land to this environmental project. However, the oaks' population will negatively affect the production of honey since bees get lower quality pollen, which ultimately increases the production costs of honey.

In the second part of this paper I allow for the possibility that every agent who decides not to participate in the conservation program can benefit from a *positive* externality from his neighbor (who is developing an environmental project). An example of positive externality in this context is the case in which the government is interested in increasing the biodiversity of a lake. Let us now assume that in this area there are two landholders, one of them is dedicated to a fish farm and the other produces vegetables that are sold in the market. If the last farmer participates in the procurement of the lake's environmental project, she will eliminate any pesticide used in the crop of vegetables and will stop dumping toxic waste to the lake. Therefore, the fish farmer has a direct benefit from the environmental project that his neighbor develops, which finally results in a decrease of his costs of fish production.

There are several authors who have analyzed externalities using mechanism design. Jehiel, Moldovanu and Stacchetti (1996) and (1998) define the bidder's valuation as a vector, formed by bidder i's valuation for winning and the valuation that he obtains when his opponents win the auction, i.e. the externality that he suffers when the object is assigned to a different bidder than him. Skreta and Figueroa (2004) also analyze the externality problem but for the case of auctions with multiple objects. The difference between the existing literature and this paper is that I consider the externality into the cost function and not in the bidder's valuation, which simplifies the solution of the optimal mechanism.

There also exists a vast literature on multidimensional messages or bids. Hansen (1988) and Desgagne (1988) were some of the authors who first considered two-dimensional bid auctions. Dasgupta and Spulber (1990) consider implementation by per unit bid auctions, where firms offer per unit bids against a given quantity schedule. Che (1993) analyzes how to implement the optimal mechanism via a score-based system of two-dimensional auctions (quality, price) and Branco (1997) analyzes the impact of costs' correlation on the design of multidimensional mechanisms. The double dimensionality in this work is represented by the message that the landholder submits in the mechanism, that is the biodiversity level that he implements and the cost associated to every level of biodiversity improvement that he receives in the form of a transfer.

The paper is structured as follows. In the next section I describe the model, where all the analysis is developed under the assumption of negative externalities. In Section 3, I identify the optimal revelation mechanism. In this part of the paper I find that the negative externality plays an important role in the specification of the participation constraint. This is due to the fact that the landholders' reservation utility is type dependent and endogenous; in other words, it depends on the cost change that a landholder experiences when his neighbor implements a particular biodiversity project. In Section 4, I analyze the consequences of having positive externalities in the mechanism, and an example is developed to understand the problem that this assumption generates in the mechanism design problem. Finally, section 5 summarizes the main results.

3.2 MODEL (WITH NEGATIVE EXTERNALITY)

The objective of the model is to analyze the dilemma that the government faces when it has to decide the optimal biodiversity to achieve in some specific area, and the cost that it needs to incur in order to implement it. Additionally, the following model studies the effects that the externality generated by the biodiversity project has on the nonparticipant landholders.

I consider a mechanism in which the government (seller) owns an indivisible good, that is the Conservation Program. There exist N risk-neutral[2] landholders indexed by $i = 1, 2, ..., n$, who are referred to as potential bidders. For simplicity, I assume that the landholders own the land that the government is interested in improving. Therefore, they are directly affected by an increase or decrease in biodiversity. Landholders submit messages of the form (B_i, t_i) for a single project that will be developed in his own land, where t_i represents the total cost of the project for the landholder, and $B_i \epsilon \mathbb{R}_+$ is a non-monetary variable that reflects the biodiversity improvement in his land due to that particular project. Similarly to Che's (1993) model, I assume that all non-monetary characteristics related to biodiversity can be aggregated into a one-dimensional variable B_i, that henceforth will be referred to as "environmental quality."

The implementation of a biodiversity project implies that the landholder has to incur some costs. The amount of costs will vary depending of the landholder's type, $\theta \in \Theta$. Assume that each landholder has private information about his costs of implementing a market activity or biodiversity project, denoted by the parameter θ. Landholders' cost parameters are independent and identically distributed with cumulative distribution of the cost parameter given by $F(\cdot)$ defined on $\left[\underline{\theta}, \overline{\theta}\right]$ which is common knowledge among both landholders and the government.

Landholder i's cost function when he participates in the conservation program is repre-

[2]Landholders are sometimes considered to be risk averse. Empirical studies assessing landholder's conservation attitudes in this respect, however, do not reach a unanimous judgement. Hence, the risk-neutrality assumption neither seems too restrictive nor contrary to empirical evidence.

sented by $C_i(B_i,\theta)$. I assume that costs are increasing in the environmental quality, $C_B > 0$, at an increasing rate, $C_{BB} > 0$, i.e. the cost function is increasing and convex in B_i. Moreover, $C_i(0,\theta) = 0$ which implies zero fix cost. Additionally, I consider that higher types of θ –more efficient landholders– have lower marginal costs of improving environmental quality, $C_{\theta B} < 0$. Furthermore, the landholder i's cost function when he produces a market good is $C_i^M(q, B_{-i}, \theta)$. It is increasing in the amount of market good produced, $C_q^M > 0$, at an increasing rate, $C_{qq}^M > 0$. Similarly to the cost function of producing biodiversity, I consider that higher types of θ have lower marginal costs of producing market goods, $C_{\theta q}^M < 0$. Below I specified how $C_i^M(q, B_{-i}, \theta)$ increases in the case of negative externalities or decreases in the case of positive externalities.

The government's utility when landholder i develops the conservation program is the following:

$$U_g(B_i, t_i) = V(B_i) - t_i$$

where t_i represents the government's transfer to landholder i. Hence, it is important to take into account that when the landholder is proposing a project that generates an environmental quality (in biodiversity terms) of B_i, he also produces a utility to the government that is represented by function $V(B_i)$. In particular, I assume that function $V(B_i)$ is increasing and concave in B_i, i.e. $V_{BB}(\cdot) < 0 < V_B(\cdot)$, with $V(0) = 0$. That is, if the project does not generate any biodiversity improvement then the government's utility from it equals zero.

Moreover, I consider that the government maximizes expected social welfare. The social welfare consists of the sum of the utility that the government derives from each landholder's implemented biodiversity level, $V(B_i)$, minus the welfare cost of raising public funds, $(1+\lambda)t_i$.

$$W = \sum_{i=1}^{n} (V(B_i) - (1+\lambda)t_i)$$

On the other hand, landholders i's utility is represented by:

$$\begin{aligned} U_l(B_i, t_i) &= \phi(i \text{ produces biodiversity})[t_i - C_i(B_i,\theta)] + \\ &\quad \phi(i \text{ does not produce biodiversity}) \left[\pi - C_i^M(q, B_{-i}, \theta)\right] \end{aligned}$$

where $\phi(\cdot)$ is an indicator function $\phi(\cdot) \in \{0,1\}$, and $\phi(\cdot) = 1$ if landholder i produces biodiversity and $\phi(\cdot) = 0$ if landholder $-i \neq i$ develops de conservation program. In the last case, I assume that landholder i dedicates his land to some market activity q where he obtains profits equal to π (which are assumed to be the same for all the non-participating landholders). However, since landholder $-i$ develops a project that increases the biodiversity, a negative externality is generated to the non-participating landholders, B_{-i}, that can be translated into an increase in landholder i's costs. Notice that in this part of the paper I assume that the environmental improvement generates a negative externality which is transmitted through an increase in the cost of all other landholders –those who do not participate in the auction– who work in a non-environmental (market) activity, q. Additionally, in spite of receiving the same environmental negative externality, B_{-i}, the effect that this externality causes in two different non-participating landholders is distinct, since each of them has a different cost function $C_i^M(q, B_{-i}, \theta_i)$ given his own realization of the efficiency parameter θ_i, $C^M_{\theta B_{-i}} > 0$. In other words, the externality is type dependent.

For the landholders who are not participating in the conservation program, their cost function, $C_i^M(q, B_{-i}, \theta_i)$, is increasing in q and, given the negative externality assumption, we have that $C^M_{qB_{-i}} > 0$, i.e. the marginal cost of producing q increases in the negative externality received B_{-i}. For simplicity, I consider that any landholder i, when receiving no externalities, $B_{-i} = 0$, is as efficient producing the market activity as he is generating biodiversity (when he wins the contract) for any realization of θ_i. Finally, in order to clarify the structure of this model, I describe the move order of the game as follows.

1. Nature chooses the type of every landholder, θ, which is distributed in $\left[\underline{\theta}, \overline{\theta}\right]$.This type conditions the landholder's cost function. The landholder privately observes his own type. He does not know the others landholders' types, and the government cannot observe the realization of any landholder's type.
2. The government wants to improve the biodiversity level of some specific area.
3. All landholders who live in the area report their messages composed by the amount of biodiversity that will be achieved and the cost of the project associated with this biodiversity level, which he will receive in the form of a transfer from the government.

4. The government chooses the optimal project to be implemented.
5. Those landholders who do not produce biodiversity dedicate their land to some market activity.
6. Those landholders who produce biodiversity generate some negative (or positive in section 4) externality to the all nonparticipating landholders in their surroundings.

3.3 OPTIMAL REVELATION MECHANISM

In this section I describe and solve the optimal direct revelation mechanism that the government uses in order to maximize the expected social welfare from the biodiversity projects, subject to the landholders' incentive compatibility conditions and their individual rationality (or voluntary participation) constraints. Generally, in this context a mechanism (p, B, t) has the following components: a message (B, t) for each buyer i; an allocation rule, p_i, which determines the probability that landholder i will get the object and a payment rule t. Notice that

$$p : \Theta \to \Sigma, \text{ where } \Sigma := \left\{ \sigma \in \mathbb{R}_+^N | \sum_{i=1}^{N} \sigma_i \leq 1 \right\}$$

is the set of probability vectors. For simplicity, I work with a direct revelation mechanism, where the strategy space coincides with the type space, and thus, each agent's strategy is just to announce a type. Hence thereafter, I use the Revelation Principle, and as a consequence the mechanism can be expressed as $(p(\theta_i), B(\theta_i), t(\theta_i))$, and understood as a mapping from the set of revealed types Θ to the set of allocation rule $p_i \in [0, 1]$, biodiversity improvement $B_i \in \mathbb{R}_+$ and transfers $t_i \in \mathbb{R}_+$.

Notice that in any mechanism function p_i must satisfy $\sum p_i(\theta_i) \leq 1$ and $p_i(\theta_i) \geq 0$. I restrict, however, the analysis to deterministic revelation mechanisms, given that the government never wants to randomize over the selected landholders. Therefore the landholders selected by the procurer implement a biodiversity improvement in their land, while all the other landholders have a project with zero biodiversity improvement. The former can be interpreted as not developing the project in their land. Then the mechanism can be expressed by a pair of functions $(B(\theta_i), t(\theta_i))$[3].

In order to ensure that the auction mechanism is feasible, it must satisfy the *participation constraints* (P.C.) and that the *incentive compatibility* (I.C.) conditions hold for all N landholders. In particular, the mechanism $(B(\theta_i), t(\theta_i))$ satisfies the participation constraint if

$$U_i(B_i(\theta_i), t_i(\theta_i), \theta_i) \geq \pi - E_\theta[C_i^M(q, B_{-i}, \theta_i)] \quad \forall \theta_i \in \left[\underline{\theta}, \overline{\theta}\right] \tag{P.C.}$$

where the reservation utility, $\pi - E_\theta[C_i^M(q, B_{-i}, \theta_i)]$, represents landholder i's net benefit when he does not participate in the auction, i.e. π is the landholder i's revenue when he dedicates his land to some economic activity, where $\pi \leq t$. Since the seller is implementing the mechanism, he can control both $B(\theta_i)$ and $t(\theta_i)$. However, he needs to construct an expectation about the expected cost of the landholder when not participating in environmental activities, given that the principal does not know the realization of θ for every landholder.

Additionally, this participation constraint is type-dependent, given that the above reservation utility is increasing in θ, i.e. higher values of θ imply lower marginal costs when dedicating his land to market activities, and as consequence higher reservation utilities. In contrast, this reservation utility is decreasing in the negative externality, B_{-i}, since an increase in B_{-i} increases the production cost of the market activity q which finally decreases the reservation utility. Hence, the rate of increase of the marginal cost of producing biodiversity, $C_{BB}(B_i, \theta_i)$, is lower than the rate of increase of the marginal cost of producing an economic good (taking into account the negative externality produced by his neighbors), $C_{qB}^M(q, B_{-i}, \theta)$ for any θ. That is,

$$C_{BB}(B_i, \theta_i) < C_{qB}^M(q, B_{-i}, \theta_i) \;\; \forall \theta_i \in \left[\underline{\theta}, \overline{\theta}\right]$$

[3]Note that in more general mechanisms B_i and t_i may depend on others landholders' types. However, this will generate a verification problem, since the transfer landholder i receives would depend on types he does not observe. In order to overcome this problem I restrict the model to those mechanisms in which B_i and t_i depend only on landholder i's own type.

In other words, the cost function of developing market activities and receiving a negative externality from your neighbors lies above the cost function of developing environmental activities –and not receiving any kind of negative externality– for all values of θ. That is, the previous assumption guarantees that all the participants in the procurement auction have incentives to participate[4].

Moreover, the mechanism $(B(\cdot), t(\cdot))$ is said to be incentive compatible for landholder i if he experiences a greater utility by revealing his type θ_i truthfully than by reporting any other type $\widehat{\theta}_i \neq \theta_i$.

$$U_i(B_i(\theta_i), t_i(\theta_i), \theta_i) \geq U_i(B_i(\widehat{\theta}_i), t_i(\widehat{\theta}_i), \theta_i) \quad \forall \theta_i \in \left[\underline{\theta}, \overline{\theta}\right] \text{ and } \forall\ \widehat{\theta}_i \neq \theta_i \qquad \text{(I.C.)}$$

Finally, let us assume that the hazard rate function $\psi(\theta) = \frac{f(\theta)}{1-F(\theta)}$ is increasing in θ. The former assumption is a sufficient condition to ensure that the design problem is regular.

Using the revelation principle, it is possible to define a direct revelation mechanism $(B(\cdot), t(\cdot))$ which can be used in the government's expected social welfare maximization problem.

$$\max_{\{B_i(\cdot), t_i(\cdot)\}} \quad E\left[\sum_{i=1}^{N} (V(B_i(\theta_i)) - (1+\lambda) t_i(\theta_i))\right]$$

Moreover, landholder i's utility of truthfully revealing his type θ_i and being assigned, as a consequence, a project with biodiversity level $B_i\left(\theta_i\right)$ and transfer $t_i\left(\theta_i\right)$ is:

$$U_i(B_i, \theta_i) = t_i(\theta_i) - C_i(B_i(\theta_i), \theta_i)$$

[4] The negative externality only affects the cost of producing a market good. In the absence of externalities the landholder can produce a market good or a biodiversity project at the same marginal cost. In the case of negative externalities, however, the growth rate of the marginal cost of producing a biodiversity project is always lower than that of producing a market good.

Therefore, the government's maximization problem can be expressed as follows:

$$\max_{\{B_i(\cdot),U(\cdot)\}} E\left[\sum_{i=1}^{N}(V(B_i(\theta_i)) - (1+\lambda)[C_i(B_i(\theta_i),\theta) + U_i(\theta_i)])\right]$$

subject to

$$B(\cdot) \text{ is nondecreasing in } \theta \in \left[\underline{\theta}, \overline{\theta}\right] \tag{1}$$

$$U_i(\theta_i) = U_i(\underline{\theta}_i) - \int_{\underline{\theta}}^{\theta_i} \frac{\partial C_i(B_i(x),x)dx}{\partial \theta_i} \; \forall \theta_i \in \left[\underline{\theta}, \overline{\theta}\right] \quad \text{(Incentive compatibility, see appendix)} \tag{2}$$

$$U_i(\theta_i) \geq (\pi - E_\theta[C_i^M(q, B_{-i}, \theta_i)]) \quad \text{(Participation constraint)} \tag{3}$$

That is, the government wants to maximize the expected social welfare generated by the biodiversity project that every landholder develops, subject to the fact that these contracts with the landholders are feasible, i.e. they satisfy the incentive compatibility and participation constraint conditions specified above. Solving the government's maximization problem (see appendix) we obtain the following proposition.

Proposition 1

Let (t^*, B^*) be the optimal mechanism. Then, the optimal menu of contracts specifying a biodiversity level to be implemented, $B_i^*(\theta)$ and a transfer $t_i^*(\theta)$ to be received for the implementation of such biodiversity level must satisfy,

$$\frac{V(B_i(\theta_i))}{dB_i} = (1+\lambda)\left(\frac{\partial C_i(B_i(\theta_i),\theta_i)}{\partial B_i} - \frac{1-F(\theta_i)}{f(\theta_i)}\frac{\partial^2 C_i(B_i(\theta_i),\theta_i)}{\partial B_i \partial \theta_i}\right)$$

and

$$t_i^*(\theta) = (1+\lambda)\left(C_i(B_i(\theta),\theta) - \frac{1-F(\theta_i)}{f(\theta_i)}\frac{\partial C_i(B_i(\theta),\theta)}{\partial \theta}\right) \quad \forall\, \theta \in \left[\underline{\theta}, \overline{\theta}\right] \text{ and } \forall\, i \in N$$

Proof. See appendix.

Hence, the optimal biodiversity that a landholder generates depends on his efficiency level, the welfare cost of raising public funds and on the hazard rate of its private parameter, θ. On the other hand, the optimal transfer is equal to the landholder's cost when he dedicates his land to develop a biodiversity project. Additionally, from the above proposition we can infer that those landholders who have low types (θ) will obtain a higher transfer if they decide to participate in the procurement. This is due to the fact that they are not so efficient in producing biodiversity, and as a consequence the government has to make a higher effort (in terms of transfers) to induce them to participate in a biodiversity project. On the other hand, efficient landholders receive lower transfers since they are more affected by the negative externality than their inefficient counterparts. This reduces their incentives to continue the production of market goods, ultimately inducing them to produce biodiversity.

3.4 BILATERAL POSITIVE AND NEGATIVE EXTERNALITY

In sections 2 and 3 we assumed that the externality suffered by the landholder was always negative. However, in this section I want to relax the former assumption. As a consequence, thereafter I assume that agents can receive negative or positive externalities from the biodiversity project that his neighbors develop[5]. Hence, the biodiversity quality generated by a landholder $-i$ and received by all of his neighbors i ($-i \neq i$), who are not participating, can be $B_{-i} \lesseqgtr 0$.

All the assumptions coincide with those in sections 2 and 3 except for the participation constraint. The reason why the participation constraint differs is explained as follows. The reservation utility is –as in the previous section – type dependent and increasing in θ, but now it is *increasing* in B_{-i}, in contrast to the previous section where it was decreasing. Indeed, now the positive externality produced by surrounding landholders decreases landholder i's costs of developing market activities, which in turn increases his reservation utility. This makes voluntary participation of all landholders more difficult to guarantee than in the case

[5] I do not analyze the case where B_{-i} is equal to zero, given that in this case the reservation utility will simply be represented by the market profits, π, which can be normalized to zero. Therefore, the mechanism design problem when $B_{-i} = 0$ becomes the classical one. See Fudenberg and Tirole (1991).

		Low Types	High Types
Positive	$C_{BB} > C_{qB}^{M}$	Participate	No incentives to participate
Negative	$C_{BB} < C_{qB}^{M}$	Participate	Participate

Table 1: Differences between positive and negative externalities

of negative externalities, since a landholder's payoff when not participating is higher in the case of positive externality than when receiving a negative one.

Given the positive and negative externality, the cost function for the landholder who decides to not participate in the conservation program can be generalized as follow,

$$C_i^M (q, B_{-i}, \theta_i) = C_i(q, B_1, ..., B_{i-1}, B_{i+1}, ..., B_N, \theta_i)$$

where landholder i is producing a market product q and he is receiving positive or negative externalities from landholders that are developing a biodiversity project. The above cost function can also be expressed as $C_i^M(q, \sum_{-i \neq i}^{N} B_{-i}, \theta_i)$. Hence, we have to impose some strict assumption on the marginal cost function since it can be increasing or decreasing depending on the value of B_{-i}, that is $C_{qB_{-i}}^M(\cdot) \gtrless 0$. Table 1 highlights the differences between the presence of negative externalities, as in the previous sections, and positive externalities in the model.

In order to understand the effects that a positive or negative externality has in landholders' cost function –and as a consequence in his decision to participate in the conservation program– we will develop the following example.

Example

Let us assume that there exist three landholders in some specific area where the government is interested in improving the biodiversity. The market benefits that these three landholders can obtain when they are dedicating their land to produce some market activity is equal to $\pi = 30$. Moreover, when a landholder does not participate his costs and the

Landholders	$\mathbf{B_1}$	$\mathbf{B_2}$	$\mathbf{B_3}$	**Costs**
Landholder 1	0	6	-10	30
Landholder 2	5	0	-6	15
Landholder 3	-7	-9	0	15

Table 2: Example of positive and negative externalities

biodiversity impact that he experiences from the participating landholders is presented in table 2.

Notice that numbers with positive sign represent positive externalities while numbers with negative signs represent negative externalities that landholder i suffers from his neighbors $-i \neq i$ when he does not produce biodiversity. For instance, when landholder 2 does not participate he receives a positive externality from the environmental project developed by landholder 1 of five units, and a negative externality of six units from landholder 3's activities. Finally, the last column expresses landholder 2's costs from continuing his market activities. Now, with all the above information we can analyze the participation constraint for each landholder.

$$\textbf{Landholder 1:} \quad U_1(B_1(\theta_1), t_1(\theta_1), \theta_1) \geq \pi - C_1^M(q, \sum_{-i \neq 1} B_{-i}, \theta_1)$$
$$\Longleftrightarrow U_1(B_1(\theta_1), t_1(\theta_1), \theta_1) \geq 30 - (30 - 6 + 10) = -4$$

$$\textbf{Landholder 2:} \; U_2(B_2(\theta_2), t_2(\theta_2), \theta_2) \geq \pi - C_2^M(q, \sum_{-i \neq 2} B_{-i}, \theta_2)$$
$$\Longleftrightarrow U_2(B_2(\theta_2), t_2(\theta_2), \theta_2 \gtreqless 30 - (15 - 5 + 6) = 14$$

$$\textbf{Landholder 3:} \; U_3(B_3(\theta_3), t_3(\theta_3), \theta_3) \geq \pi - C_3^M(q, \sum_{-i \neq 3} B_{-i}, \theta_3)$$
$$\Longleftrightarrow U_3(B_3(\theta_3), t_3(\theta_3), \theta_3) \geq 30 - (15 + 7 + 9) = -1$$

Clearly landholders 1 and 3 have incentives to participate in the biodiversity program. Landholder 2, however, has incentives to be a free rider –given the positive externality that he receives from his neighbor– and to not participate in the procurement. In particular, the government would have to offer him some transfer greater than his market net benefit of 14 to induce him to participate. Additionally, note that landholders 2 and 3 have exactly the same cost (we can say that they are equally efficient in market activities). However, landholder 3 is receiving a very high negative impact when not participating, what finally induces him to participate in the auction, since he will obtain a higher utility developing an environmental project than conducting a market activity. Therefore, from the example we can infer that the impact of a positive externality may generate incentives to not participate in the environmental project for some landholders, i.e. landholder 2. ▲

Let us focus from now on in the case where the landholder is enjoying an overall positive impact from externalities $\sum_{-i=1}^{N} B_{-i} > 0$. The participation constraint was defined as,

$$U_i(B_i(\theta_i), t_i(\theta_i), \theta_i) \geq \pi - E_\theta[C_i^M(q, B_{-i}, \theta_i)]$$

In the case of positive externalities $C_{qB}^M < 0$, i.e., the marginal cost of developing market activities for the nonparticipating landholder decreases in the positive externality he receives from the participating neighbors. Additionally, the growth rate of the marginal cost of participating in an environmental project, C_{BB}, is higher or equal than the growth rate of the marginal costs of production in the market, in absolute value, C_{qB}^M, i.e., $C_{BB} \geq C_{qB}^M$. Obviously in this situation those types with a lower θ will be interested in participating, but not those with higher θ.

The optimal transfer obtained in proposition 1 is not applicable for the case of positive externalities. As commented above, this is due to the fact that the reservation utility specified for the participation constraint is nondecreasing in θ –more efficient landholders find market activities more convenient than less efficient landholders– which induces these efficient landholders to not participate. Therefore, given the above restriction, I introduce a useful assumption that will help us find an optimal solution when the non-participating landholders are enjoying a positive externality from their neighbors. Specifically, from now on, I assume that the government is interested in developing the biodiversity project in those areas in which landholders are less efficient, i.e,. low θ.

The above assumption can be rationalized by considering the following example. City councils usually develop projects, such as the construction of a museum or a public university, in an area that is considered a "bad" neighborhood. In spite of this adjective, however, the city council expects that this project can positively transform the neighborhood where it is being implemented. In our particular case, the government is interested in developing the biodiversity project in those lands where the landholder is less efficient (low θ). Given that higher types are more efficient producing the market activity, q, they will obtain more benefits in the market than those landholders who have lower types. So, we need to find a cutoff that induces the participation of those inefficient types who belong to some range $\left[\underline{\theta}, \widehat{\theta}\right]$, where $\widehat{\theta} < \overline{\theta}$, which I find in the next proposition. In addition, from section 3, we know that the first order condition of the relaxed problem is given by the following expression, which I use next.

$$\frac{V(B_i(\theta))}{dB_i} - (1+\lambda)\frac{\partial C_i(B_i(\theta), \theta)}{\partial B_i} + (1+\lambda)\frac{\partial^2 C_i(B_i(\theta), \theta)}{\partial B_i \partial \theta_i}\frac{1-F(\theta_i)}{f(\theta_i)} = 0 \;\forall \theta \in \left[\underline{\theta}, \overline{\theta}\right] \tag{4}$$

Proposition 2

In the case of a positive externality, the optimal transfer $t^*(\theta)$ is given by

$$t_i^*(\theta) = (1+\lambda)\left(C_i(B_i(\theta), \theta) - \frac{1-F(\theta_i)}{f(\theta_i)}\frac{\partial C_i(B_i(\theta), \theta)}{\partial \theta}\right)$$

for all $\theta \in \left[\underline{\theta}, \widehat{\theta}\right]$, where $\widehat{\theta} < \overline{\theta}$. In particular, $\widehat{\theta}$ is the type for which

$$V(B_i(\widehat{\theta})) = (1+\lambda)\left(C_i(B_i(\widehat{\theta}), \widehat{\theta}) - C_i^M(q, B_{-i}, \widehat{\theta})\frac{1-F(\widehat{\theta}_i)}{f(\widehat{\theta}_i)} - \frac{\partial C_i(B_i(\widehat{\theta}), \widehat{\theta})}{\partial \widehat{\theta}}\frac{1-F(\widehat{\theta}_i)}{f(\widehat{\theta}_i)}\right)$$

and all landholders with types $\theta \in (\widehat{\theta}, \overline{\theta}]$ do not participate in the auction.

Note that for the highest type $\left(\overline{\theta}\right)$ the above result is obvious. If we assume that $\theta \sim U[0,1]$, and using expression (4) for $\overline{\theta}$, the inverse hazard rate becomes zero and therefore the optimal transfer for the highest type is equal to the cost of the project. However, given the assumption $C_{BB} \geq C_{qB}^M$, this implies that $\overline{\theta}$ –the most efficient landholder– does not participate in the procurement. On the other hand, for all landholders with lower types $\theta \in \left[\underline{\theta}, \widehat{\theta}\right]$, where $\widehat{\theta} < \overline{\theta}$, we have that the marginal cost of producing a market activity is equal to the marginal cost of generating biodiversity given a level of received (respectively, produced) externality in the interval $\left[0, \widehat{B}\right]$. That is,

$$C_q^M\left(q, B_{-i}, \theta_i\right) = C_B\left(B, \theta_i\right)\ \forall \theta_i \in \left[\underline{\theta}, \widehat{\theta}\right]$$

and the marginal cost of developing a market activity is lower than the marginal cost of producing biodiversity for biodiversity levels in the interval $\left(\widehat{B}, +\infty\right)$, as the following figure 3 illustrates.

$$C_q^M\left(q, B_{-i}, \theta_i\right) < C_B\left(B, \theta_i\right)\ \forall \theta \in \left[\widehat{\theta}, \overline{\theta}\right]$$

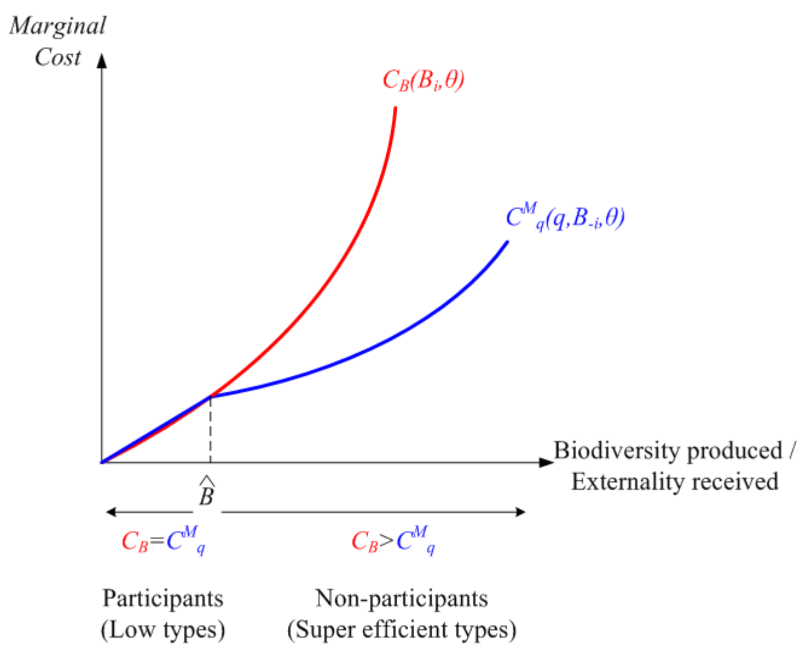

Figure 3: Low types produce biodiversity

3.5 CONCLUSIONS AND EXTENSIONS

The existence of negative or positive externalities in the development of a biodiversity project has clear consequences in the solution of an optimal mechanism that solves the government's expected social welfare maximization problem. Specifically, the assumption that the externality directly affects the cost function of the nonparticipating landholders in the surroundings gives us a new perspective to analyze the consequences of externalities in the design of a procurement auction.

In the case of a negative externality I found that the optimal transfer depends on the efficiency of the landholder. In particular, this paper shows that those efficient landholders with high types receive lower transfers. This result is based on the assumption that the rate of increase of the marginal cost of producing biodiversity, $C_{BB}(B_i, \theta)$, is lower than the rate of increase of the marginal cost of producing an economic good (taking into account the negative externality received from his neighbors), $C^M_{qB}(q, B_{-i}, \theta)$, what induces all landholders to participate in the conservation program.

On the other hand, positive externalities seem to be more involved than negative ones. Specifically, the landholder's reservation utility –and specified in the participation constraint of the government's problem– is nondecreasing in his type, θ. This implies that higher types are not willing to participate in the development of a biodiversity project, since positive externalities decrease the marginal cost of developing market activities for those landholders who do not participate in the environmental project, i.e., $C^M_q(q, B_{-i}, \theta) < C^M_B(B_i, \theta)$.

In order to deal with positive externalities, I assumed that the government is only interested in inducing the participation of those less efficient landholders. Therefore, those landholders who live in the area where the government wants to improve the biodiversity and are less efficient (low types) will be the ones selected to develop the biodiversity project. In contrast, the more efficient landholders of such area will enjoy the positive externality that will result in a decrease of their cost of producing market activities. In addition, as in the negative externality case, an optimal biodiversity schedule $B_i^*(\theta)$ and an optimal transfer function $t_i^*(\theta)$ are identified in the presence of positive externalities.

The contributions of the paper can, then, be analyzed with respect to the mechanism design literature on one hand, and the environmental economics efforts on methods to evaluate environmental quality on the other hand. Firstly, regarding the mechanism design literature, this paper considers the possibility of positive and negative externalities among players by introducing them into the landholders' cost functions, what greatly simplifies the analysis –as compared to Jehiel, Moldovanu and Stacchetti (1999). Secondly, with respect to the environmental economics literature dealing with traditional methods of evaluating the environmental impact of a project, this paper shows that we can design mechanisms where participants are induced to truthfully report the effects they receive from an externality. Indeed, most of the methods proposed by this literature –such as the Cost-Benefit analysis and the Cost Effectiveness method– are constantly criticized because of their inability to induce a truthful revelation of the effect that a certain externality may have in the utility function of those individuals affected by the externality. Through the mechanism described in this paper, however, this information problem is solved by the use of an appropriately designed procurement auction.

Further extensions can be considered for this paper. For example it may be interesting to analyze the model developed in this study but considering risk averse landholders. This assumption is more realistic, given that usually landholders are affected by uncontrollable variables such as weather, natural disasters, etc., which might induce a more risk averse behavior. The correlation of the landholders' types or assuming the externality as a random variable, uncontrollable by the government, can both be considered for further research.

4.0 THE IMPORTANCE OF FOREGONE OPTIONS: GENERALIZING SOCIAL COMPARISONS IN SEQUENTIAL-MOVE GAMES

4.1 INTRODUCTION

Recent advances in behavioral economics allow for the possibility that individuals care about the payoffs of others. In particular, most of these advances suggest the existence of *social*, as opposed to *individual*, preferences reflecting individuals' predilection for fairness in the income distribution; see Fehr and Schmidt (1999) and Bolton and Ockenfels (2000).

Despite the multiple situations that can be rationalized with these approaches, a recent literature suggests that individuals' behavior cannot be explained by theories on social preferences alone. Specifically, an agent's choices can only be supported by analyzing how she evaluates other players' chosen and *unchosen* actions. For example, Brandts and Solà (2001), Falk *et al.* (2003) and Charness and Rabin (2002) accumulate significant evidence supporting the importance of unchosen alternatives in the ultimatum bargaining game, while Andreoni, Brown and Vesterlund (2002) show the relevance of unchosen alternatives in public good games. In order to illustrate their results, let us briefly analyze Brandts and Solà's (2001) study. In particular, they examine an ultimatum bargaining game in which the proposer is called to choose among only two alternative divisions of the pie (which we normalize to a size of one) as the figures 4 and 5 below illustrate.

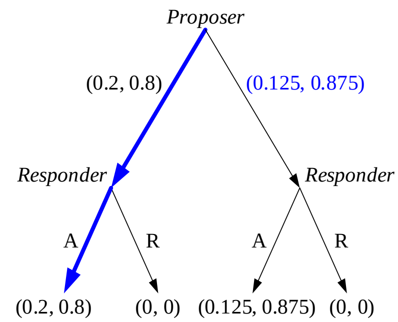

Figure 4: Responder accepts

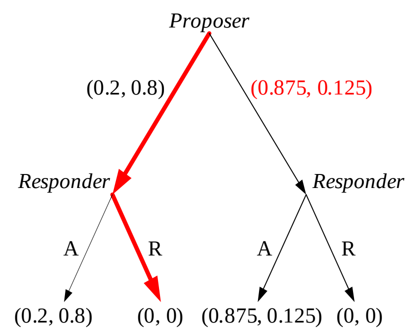

Figure 5: Responder rejects

Specifically, they consider two treatments. In the first one, represented in figure 4, the proposer chooses among two divisions of the pie, (0.2,0.8) and (0.125,0.875) —where the first and second component of every pair denote the receiver and proposer's payoff, respectively. In the second treatment, as figure 5 indicates, the first available division (0.2,0.8) is unchanged, while the second division becomes (0.875,0.125). Importantly, they show that, conditional on division 0.2 being offered to the responder (bold lines in the figures), the proportion of receiver's rejections is significantly higher when the unchosen division of the pie that the proposer *did not* select was 0.875 (figure 5) than when it was 0.125 (figure 4). That is, for a given offer 0.2 to the responder, the proportion of rejections increases in the share of the pie that the responder could have received. Intuitively, the receiver positively evaluates a given offer when the alternative division of the pie is below the actual offer that the proposer makes him (he infers "kindness"), and negatively otherwise (he interprets "unkindness"). Certainly, the receiver's pattern of rejections cannot be rationalized using inequity aversion. Indeed, once offer (0.2,0.8) is made in both treatments, the inequity in the payoff distribution is *constant* across treatments, and yet the receiver's behavior is *different* across treatments. The receiver's rejecting pattern cannot be explained using chosen actions either, since the proposer's chosen offer is constant across treatments but the receiver's behavior is not. Instead, any rationalization of the previous results must rely on the receiver's comparison between the actions that the proposer chooses and those he does not (*unchosen* actions).

References to unchosen actions are nevertheless not restricted to economic contexts alone. For instance, we frequently encounter references to unchosen alternatives in the way in which many national and international policies are announced to the media. Indeed, these public presentations are often accompanied with statements like "The government/organization/firm had to choose between policies A and B, and choosing A would have been so much worse that we finally decided to select B." These statements are certainly effective when they induce the listener to positively evaluate the chosen action B *relative to* the unchosen action A.

In this study we introduce a model that rationalizes these economic conducts in complete information sequential-move games within a general framework of economic behavior. Specifically, we assume that as in standard models, every player cares about her material payoff. Additionally, we consider that every individual compares other players' actually chosen actions with respect to a particular action that they could have selected (other players' foregone actions). Hence, this particular action is used by every individual as a *reference point* to measure the kindness she perceives from other players' choices.

This paper makes two main contributions. *First*, we identify conditions under which players' equilibrium actions are higher when individuals are concerned about these reference-dependent comparisons than when they are not. In particular, this set of conditions allow for a direct prediction about whether players cooperation when they are concerned about relative comparisons is either higher or lower than in standard game-theoretic models, for a broad class of relative comparisons players might use. Importantly, it examines players' cooperation even when they are not concerned about each other's material payoffs. Indeed, unlike models with inequity averse individuals where players do care about other individuals' payoffs (social preferences), this paper analyzes conditions under which agents choose higher strategy levels than in standard models without the need to assume that they care about other people's payoffs, i.e., even when agents' preferences can be regarded as "strictly individualistic." *Second*, we show that the model this paper describes embeds as special cases existing behavioral models: from models on inequity aversion to those analyzing social status acquisition. Finally, we apply our model to different economic applications where we enhance the intuitions behind our results: the ultimatum bargaining game, the labor market gift exchange game, and the sequential public good game. Our equilibrium predictions are not only validated in these applied settings, but also confirmed by recent experimental data.

The structure of the paper is as follows. In the next section we discuss the literature on social preferences and intentions-based reciprocity, their relationship with our paper, and how it complements their approach. In section three, we describe the properties that players' utility function must satisfy in order to support our results in terms of higher degrees of cooperation. Furthermore, section four analyzes players' equilibrium strategy in these sequential-move games, and section five applies the model to three economic examples. Finally, the last section discusses some conclusions of the paper as well as its further extensions.

4.2 RELATED LITERATURE

4.2.1 Theoretical literature on social preferences

The literature on behavioral economics has extensively considered elements other than one's own payoff in individuals' utility function. This literature mainly deals with the so-called "other regarding preferences," since most of the papers in this area focus their attention on analyzing to what extent players care about the payoffs of his competitors, or about the distribution of payoffs in the entire population. In this respect, some papers on inequity aversion, such as Fehr and Schmidt (1999) and Bolton and Ockenfels (2000) play a prominent role. On one hand, Fehr and Schmidt (1999) consider in their two-player version the following utility function for player i

$$U_i(x_i, x_j) = x_i - \alpha_i \max\{x_j - x_i, 0\} - \beta_i \max\{x_i - x_j, 0\} \tag{4.1}$$

where x_i is player i's payoff. Intuitively, α_i represents the disutility from allocations that are disadvantageously unequal for player i (i.e., he may feel envy about player j's payoffs), while β_i denotes the guilt feeling from being the agent with the highest payoff of the population.[1] Bolton and Ockenfels (2000) also develop a similar (yet more general) model of inequity aversion in which individuals' utility is assumed to be increasing and concave in their share of total income, i.e., people experience a positive but diminishing marginal utility from receiving a higher share of the total amount of social payoffs. These models of social preferences, however, cannot rationalize the puzzling experimental evidence presented in the introduction.[2] Indeed, any model which explains such results must necessarily complement the above specification by introducing the importance of unchosen alternatives into player i's utility function, as this paper examines.[3]

[1] Interestingly, Blanco *et al* (2007) present experimental evidence supporting inequity aversion at the *aggregate* level (across all participants of a particular game) but refuting it at the *individual* level (for a particular participant across games). Their results can be confirmed by our model, whereby participants of a particular game exhibit concerns for unchosen alternatives, but they may use different foregone options across games as a reference point for comparison.

[2] Another interesting experimental paper that also tests whether payoff distributions suffice to explain players' behavior in the ultimatum bargaining game is Bereby-Meyer and Niederle (2005). Specifically, they show that the responder is more likely to reject low offers when a rejection payoff is accrued to a third player —with no strategic role in the ultimatum bargaining game— than when such payoff is accrued to the proposer.

[3] Some axiomatic approaches on inequity aversion, such as Segal and Sobel (1999), examine what conditions on players' preferences must be satisfied in order to obtain utility functions which can be represented as a weighted average of a player's own material payoff as well as that of others. Despite their interest, our approach differs from theirs, since we not only include players' actually *chosen* actions in their utility function (as they do), but also players' *unchosen* actions.

4.2.2 Models on intentions-based reciprocity

As suggested above, this paper is more in the line of Charness and Rabin (2002), whereby they analyze the intentions that players express with their actual choices along the game. In particular, they assume that agents evaluate multiple characteristics of the equilibrium allocation —including fairness and intentions— by establishing different comparisons between own and social payoffs (i.e., between x_i and x_j). Specifically, when only intentions are considered, agent i's utility function in Charness and Rabin's (2002) model reduces to

$$U_i(x_i, x_j) = \begin{cases} x_i + \theta(x_i - x_j) & \text{if player } j \text{ misbehaved} \\ x_i & \text{otherwise} \end{cases} \tag{4.2}$$

where player j's misbehavior can implicitly include player i's concern about player j's foregone options, and where θ represents the importance of intentions-based reciprocity for player i. Note, however, that player i's disutility from player j's misbehavior is scaled up by the difference between player i and j's payoffs, $x_i - x_j$. Certainly, this confounds the elements triggering such perception of misbehavior (which implicitly includes unchosen alternatives), and how this misbehavior is then measured (by considering inequity aversion). Likewise, most of the experimental literature testing reciprocating behaviors triggered by kind intentions also considers that agent i measures player j's intentions by comparing x_i and x_j; see Cox (2001, 2003).

Similarly, Falk and Fischbacher (2006) recently analyze how a given player i evaluates the kindness inferred from player j's actions by also comparing their payoffs. In particular, that study measures kindness by considering the product of two elements: the above interpersonal payoff comparison (what they refer as the "outcome term"), and a measure of other players' intentions which reflects the set of available choices for these players (the "intentions factor"). Hence, Falk and Fischbacher (2006) assume that the reference standard with which players compare their own payoff is that of other players (i.e., the payoff distribution), and then they scale up this payoff distribution according to the degree of freedom in the other players' available choices.

Finally, Cox, Friedman and Sadiraj (2007) construct a nonparametric model in which a player's preferences become more altruistic with respect to other players when she infers that these players have behaved generously with her. However, their notion of generosity is not equivalent to our definition of kindness, nor their notion of altruism coincides with our definition of reciprocity, since they assume that players compare their payoffs with that of others in their group. Unlike these models, we do not introduce other people's payoffs into player i's evaluation of intentions or kindness. Instead, in our model player i measures the kindness in player j's actions by comparing player j's chosen and unchosen (foregone) actions. In the following section we describe how this comparison is made, and how it encompasses models on inequity aversion and intentions-based reciprocity as special cases.

4.3 MODEL

Let us consider complete information sequential-move games with two players and two stages. In particular, we examine games $G = \langle S_i, S_j; u_i, u_j \rangle$, in which a female leader (player j) selects an action $s_j \in S_j \subset \mathbb{R}_+$, and afterwards a male follower (player i) chooses an action $s_i \in S_i \subset \mathbb{R}_+$. The leader's action may represent, for instance, her wage offer to a worker, or her monetary contribution to a public good. Similarly, the follower's action may denote, respectively, his effort level in a labor market game, or his monetary donation to a charity in the sequential public good game. (Note that for simplicity we describe our model for continuous action spaces. Nonetheless, all our assumptions can be extended to discrete action spaces as well). Every action profile $s = (s_i, s_j) \in S_i \times S_j$ is then mapped into the set of possible outcomes by function $out : S_i \times S_j \to X$. Note that an outcome, $out(s)$, in the ultimatum bargaining game is a monetary amount, while in public good games is a pair composed of an amount of private goods and the total contributions to the public good. Finally, every player i assigns a utility value to every outcome through her utility function $u_i : X \to \mathbb{R}$.

Note that the outcome function maps every action profile into a single outcome, i.e., there is a unique action profile leading to every terminal node of the game. Hence, for every outcome $out(s) \in X$ we can identify the unique action profile $s = (s_i, s_j)$ which induces this outcome. This allows utility function $u_i : X \rightarrow \mathbb{R}$ to be represented over action profiles in the form $U_i^{NC} : S_i \times S_j \rightarrow \mathbb{R}$, i.e., $U_i^{NC}(s_i, s_j) \in \mathbb{R}$. Specifically, superscript NC denotes that player i is "not concerned" about player j's unchosen alternatives, as opposed to superscript C, which we use in the next section to refer to players who are "concerned" about each others' unchosen actions. Finally, let us henceforth denote by single (double) subscripts in the utility function its first (and second) order derivatives.

Assumption A1. Positive but decreasing marginal benefit from other players' actions, s_j. That is, $U_{s_j}^{NC}(s_i, s_j) \geq 0 \geq U_{s_j s_j}^{NC}(s_i, s_j)$ for all s_i and s_j.

Thus, every player i benefits from increases in other players' actions, but at a decreasing rate. Note that we are deliberately vague about how $U_i^{NC}(s_i, s_j)$ increases (or decreases) in her *own* action s_i. In this way, we can capture models where players' marginal utility from increasing her action is positive (e.g., contributions in public good games) as well as negative (e.g., effort in labor market games). Next, we assume that player i's utility function is strictly concave in his own actions, s_i.

Assumption A2. *Concavity.* $U_{s_i s_i}^{NC}(s_i, s_j) < 0$ for all s_i and s_j.

Note that concavity did not hold in the motivating example discussed in the introduction since players' action space was discrete and binary. Nonetheless, we introduce this assumption given that it guarantees the existence of a unique equilibrium when players' action space is continuous. In particular, uniqueness will facilitate the comparison of the equilibrium prediction when players are not concerned about unchosen alternatives, and that when players are concerned.[4]

[4] In the case of discrete and binary action spaces, as those in the motivating example of the ultimatum bargaining game, concavity is not necessary. Instead, in order to facilitate the comparison of our results and those of standard models, we only need the subgame perfect equilibrium to be unique, both when players are concerned about foregone options and when they are not.

Assumption A3. *Strategic Substitutability.* Player j's (first mover) utility function satisfies $U^{NC}_{s_j s_i}(s_i, s_j) < 0$ for all s_i and s_j.

Thus, the first mover's marginal benefit from increasing her own action, s_j, *decreases* when the second mover raises her action, s_i. That is, the leader considers the follower's actions as strategic substitutes of her own. This assumption is sensible for a large class of games, where players try to free-ride each others' actions, e.g., the first mover's incentives to free-ride the second mover's donations to the public good or his effort decision. Therefore, A3 eliminates payoff structures such as those in the impunity game, whereby (in a variation of the ultimatum bargaining game) the first mover obtains exactly the same payoff regardless of the second mover's actions, i.e., unconditional on his acceptance or rejection of the first mover's offer. In contrast, A3 maintains the first mover's incentives to free-ride the second mover's action, since she considers players' actions as strategic substitutes.

4.3.1 How kindness enters into players' preferences

As suggested in the motivating example from Brandts and Solà (2001), players' observed behavior is clearly inconsistent across the games in their example. The games they consider are nevertheless relatively similar, since only the set of available choices for the proposer is modified. In particular, we want to describe a single utility function which is general enough to be applicable to games maintaining "similar" properties, as the two treatments considered by Brandts and Solà (2001). Specifically, in this paper we regard games as being similar when the utility that player i obtains from every action profile s coincides across the games for which this action profile induces the same outcome $out(s)$, and $out(s) \in X$. (In the previous example of the ultimatum bargaining game, if a given action profile induces the same outcome across different games then the utility that players obtain from this action profile coincide across these games.) In particular, let $U^C_i(s_i, s_j)$ represent the utility function we apply to this class of games. Specifically, $U^C_i(s_i, s_j)$ is player i's utility function when he uses player j's foregone options as a measure of the kindness behind her actions. Let us first describe how this kindness enters into player i's utility function, and then analyze how players measure the kindness behind their opponent's actions.

Assumption A4. *Kindness.* For any actions $s_i \in S_i$ and $s_j \in S_j$, player i's utility function satisfies

$$\begin{aligned} U_i^C(s_i, s_j) &\geq U_i^{NC}(s_i, s_j) \text{ if kindness} \\ U_i^C(s_i, s_j) &< U_i^{NC}(s_i, s_j) \text{ if unkindness} \end{aligned}$$

Therefore, this assumption determines when player i is concerned about social comparisons and he interprets kindness from player j's actions, his utility level is higher than when he is not concerned about these comparisons. Otherwise (when he infers unkindness), his utility level is lower. Let us next describe how this kindness affects player i's marginal utility.

Assumption A5. *Reciprocity.* For any actions $s_i \in S_i$ and $s_j \in S_j$, player i's utility function satisfies

$$\begin{aligned} U_{s_i}^C(s_i, s_j) &\geq U_{s_i}^{NC}(s_i, s_j) \text{ if kindness} \\ U_{s_i}^C(s_i, s_j) &< U_{s_i}^{NC}(s_i, s_j) \text{ if unkindness} \end{aligned}$$

Hence, A5 specifies that when player i interprets kindness from player j's actions, his marginal utility from increasing s_i when he is concerned about foregone options is weakly higher than when he is not. Otherwise, his marginal utility is lower. This property is illustrated in figure 20 (see appendix). In particular, this assumption leads player i to increase his action when he infers kindness (positive reciprocity), and to decrease it when he infers unkindness (negative reciprocity).

4.3.2 How players measure kindness

Let us now describe how players evaluate the kindness behind other players' actions. In particular, we assume that player i measures kindness through the following distance function, $D_i(s_i, s_j)$, and that he infers kindness when the outcome of this distance function is positive, and unkindness otherwise.

$$D_i(s_i, s_j) = \alpha_i \left[s_j - s_j^{R_i} (s_i, s_j) \right] \tag{4.3}$$

for any $\alpha_i \in \mathbb{R}$, where α_i can be both positive or negative. Thus, player i evaluates player j's kindness by comparing player j's actually chosen action, s_j, and a particular reference action that player i uses for comparison, $s_j^{R_i}(s_i, s_j) \in S_j$, among player j's available choices, as defined below.[5] For simplicity, this distance function was chosen to be linear. Nonetheless, from a more general perspective, player i's distance function could be nonlinear, as long as it *increases* in player j's actually chosen strategy, s_j, and *decreases* in the reference action that player i uses for comparison.

We consider that this reference-dependent measure is a natural way for player i to assess player j's actions, which is yet general enough to embed different behavioral models as special cases, as this paper shows. In particular, this distance function is similar to that in the literature on reference-dependent preferences, such as Köszegi and Rabin (2006). However, their model analyzes individual decision making, unlike this paper where we examine the strategic effects of such reference-dependent preferences. On the other hand, our distance function differs from that in Rabin (1993) for simultaneous-move games and that in Dufwenberg and Kirchsteiger (2004) for sequential-move games. Indeed, these studies assume that player i compares his actual payoff with respect to the "equitable" payoff (his equitable share in the Pareto-efficient payoffs). In contrast, we allow player i to compare player j's actually chosen action with respect to *any* feasible action, $s_j^{R_i}(s_i, s_j) \in S_j$, leading to equitable or

[5] Note that, for simplicity, we assume that player i compares player j's actions, instead of the payoffs resulting from these action choices. Choosing the latter, however, would not modify our results, since player i's payoffs are increasing in player j's action choices (assumption A1). Hence, both a definition of kindness based on the payoffs that player i obtains from player j's choices and a definition directly based on these choices increase in player j's action choices.

non-equitable payoffs. This greater generality in the reference point that player i uses for comparison provides a rationalization for the experimental evidence presented in the introduction. Let us next define the concept of reference action, $s_j^{R_i}(s_i, s_j)$, which player i uses as a reference point in order to evaluate the kindness that he perceives from player j's actually chosen action, s_j.

Definition 1. *Player i's reference point function $s_j^{R_i} : S_i \times S_j \to S_j$, maps the pair (s_i, s_j) of both players' actually chosen actions, into a reference action $s_j^{R_i} \in S_j$ from player j's set of available choices. In addition, $s_j^{R_i}(s_i, s_j)$ is weakly increasing in s_i and s_j, and twice continuously differentiable in s_i and s_j.*

Hence, player i can use any of player j's available actions in S_j as a reference point.[6] That is, $s_j^{R_i}(s_i, s_j)$ is allowed to be above/below/equal to player j's actually chosen action, s_j, which leads to negative/positive/null distances, respectively. Obviously, the particular sign of such distance affects player i's utility function, $U_i^C(s_i, s_j)$, as described above. Additionally, note that when both players' strategy spaces are identical, $S_i = S_j = S$, player i's reference point function becomes $s_j^{R_i} : S^2 \to S$. In this context, the reference point function can be, for instance, $s_j^{R_i}(s_i, s_j) = s_i$ for all s_j. In such case, $D_i(s_i, s_j) = \alpha_i [s_j - s_i]$, and player i compares player j's chosen action, s_j, with respect to her own, s_i. In particular, note two specific examples of this distance function. First, when $\alpha_i > 0$, it may represent the case that $s_j > s_i$ is interpreted by player i as a signal of player j's kindness (e.g., her commitment to contribute high donations to the public good), whereas $s_j < s_i$ is evaluated by player i as a sign of unkindness by her opponent (e.g., free-riding). The second example is related to players' concerns for status acquisition. Particularly, when $\alpha_i < 0$, player i makes the same comparison, but introduces the outcome of $D_i(s_i, s_j)$ into her utility function *negatively*, i.e. $D_i(s_i, s_j) = -\alpha_i [s_j - s_i] = \alpha_i [s_i - s_j]$ In these cases, player i may evaluate $s_j > s_i$ negatively because the action space represents for example the consumption of a given positional good that enhances social status, and that player i wants to acquire.

[6] For simplicity, we restrict the range of reference points to player j's available choices, S_j. More generally, $s_j^{R_i}(s_i, s_j)$ could take values outside S_j. We believe, however, that it is more natural to assume that player i compares player j's actions with respect to her foregone options than to actions which were not even available to her.

Furthermore, we allow player i to modify the reference action he uses to compare player j's actually chosen action, i.e., $s_j^{R_i}(s_i, s_j)$ is not restricted to be constant for all s_j. In particular, we only assume that, for a given increase in player j's action, s_j, the reference point that player i uses, $s_j^{R_i}(s_i, s_j)$, does not increase as fast as player j's action, i.e., $1 \geq \partial s_j^{R_i}(s_i, s_j)/\partial s_j$. Intuitively, this condition makes higher values of player j's action meaningful for player i, since they increase the outcome of his distance function, i.e., $\partial D_i(s_i, s_j)/\partial s_j = 1 - \partial s_j^{R_i}(s_i, s_j)/\partial s_j$; and as we described above, positive distances ultimately raise player i's utility level (kindness). As a remark, note that $D_i(s_i, s_j)$ does not depend on any possible randomness over payoffs. Indeed, player i's utility level does not depend on the difference between payoffs he could have received from the outcomes of a certain lottery, but only on payoffs he could have obtained from alternative choices of the other players. This distinction differentiates our approach from regret theory, as in Loomes and Sugden (1982), since our model focuses on agent i's evaluation of other players' chosen and unchosen actions as a measure of their kindness. Finally, extending assumption A2 to the context of concerned players, we assume that $U_i^C(s_i, s_j)$ is also strictly concave in all player i's action, s_i.

4.3.3 Best response function

The previous section described the structure behind players' preferences, how they evaluate the kindness behind other players' actions, and how this kindness enters into their utility function. In this section, we turn to examine players' best response function in these games. Let $s_i^C(s_j) \in \arg\max_{s_i} U_i^C(s_i, s_j)$ denote player i's best response function when he assigns a positive importance to player j's foregone options, and $s_i^{NC}(s_j) \in \arg\max_{s_i} U_i^{NC}(s_i, s_j)$ his best response function when he does not. Let us next analyze the slope of player i's best response function.

Lemma 1. *The slope of player i's best response function when he is concerned about foregone options, $s_i^C(s_j)$, is higher than that when he is not, $s_i^{NC}(s_j)$. That is,*

$$\frac{\partial s_i^C(s_j)}{\partial s_j} \geq \frac{\partial s_i^{NC}(s_j)}{\partial s_j} \text{ for any } s_j \in S_j \tag{4.4}$$

That is, when player i assigns a positive importance to foregone options he is more *sensitive* to increases in player j's actions than when he does not. In addition to being more sensitive, the next proposition shows that in fact he actually responds more (less) cooperatively when he perceives kindness (unkindness) from player j's actions compared to how he would react in the case of being unconcerned about player j's unchosen alternatives.

Proposition 1. *Player i's best response function when he is concerned about foregone options is higher than that when he is not if player i infers kindness from player j's actions; and lower if he infers unkindness. That is,*

$$\begin{aligned} s_i^C(s_j) &\geq s_i^{NC}(s_j) \text{ for all } s_j \text{ such that } D_i(s_i, s_j) \geq 0 \\ s_i^C(s_j) &< s_i^{NC}(s_j) \text{ for all } s_j \text{ such that } D_i(s_i, s_j) < 0 \end{aligned}$$

Intuitively, player i (when concerned about player j's foregone options) responds more cooperatively to what he perceives as kind actions, $D_i(s_i, s_j) \geq 0$, than when he is unconcerned, i.e., $s_i^C(s_j) > s_i^{NC}(s_j)$. The opposite happens when he interprets that player j's actions are unkind, i.e., $s_i^C(s_j) < s_i^{NC}(s_j)$. In other words, his interpretation of kind (or unkind) actions triggers a higher (lower) response when he is concerned about foregone options than when he is not. For example, the worker in the labor market gift exchange game, when perceiving kind actions from the firm manager, exerts a higher effort when he is concerned about the firm manager's unchosen alternatives (foregone wage offers) than when he is not, and a lower effort otherwise.

4.4 EQUILIBRIUM ANALYSIS

Recall that player j represents the first mover in this complete information sequential-move game, and player i denotes the second mover. Note that player i's best response function, $s_i^C(s_j)$, in the subgame perfect equilibrium of this game was already described in the above lemma 1 and proposition 1. Let us now analyze player j's (first mover) equilibrium action in this sequential game.

Lemma 2. *The leader's marginal utility from increasing her own action s_j is higher when the follower is concerned about her unchosen alternatives than when he is not. That is, for any action $s_j \in S_j$ player j's (first mover) utility function satisfies,*

$$\frac{\partial U_j^{NC}\left(s_i^C\left(s_j\right), s_j\right)}{\partial s_j} \geq \frac{\partial U_j^{NC}\left(s_i^{NC}\left(s_j\right), s_j\right)}{\partial s_j} \tag{4.5}$$

From this lemma, the following proposition is immediately derived.

Proposition 2. *If assumptions A1-A5 are satisfied, then $s_j^C \geq s_j^{NC}$. That is, the leader's equilibrium strategy when dealing with a follower who is concerned about foregone options, s_j^C, is weakly higher than her equilibrium strategy when facing a follower not concerned about foregone options, s_j^{NC}.*

Hence, in the subgame perfect Nash equilibrium strategy profile of the game with positive concerns for foregone options the leader chooses a higher equilibrium action than that in the game with no concerns for unchosen alternatives.[7] This result is especially relevant for certain games, such as the labor market gift exchange and the sequential public good game, where the introduction of concerns for foregone options leads to higher levels of cooperation among the players. In particular, as we show in section 5 for different economic applications, the fact that the follower is sensitive to the leader's unchosen alternatives attenuates the leader's incentives to shift most of the burden to the follower (reducing free-riding) which ultimately triggers higher actions from her than in standard game-theoretic models.[8] Furthermore, the profile of actions that players choose in equilibrium, as we also show in section 5, can better rationalize experimental results of players' observed behavior.

[7] As a remark, note that the follower moves his action choice in the *opposite* direction than the first mover moves her when he regards actions as strategic substitutes (negatively sloped best response function); whereas he moves it in the *same* direction when actions are strategic complements (positively sloped best response function).

[8] These results can be easily generalized to sequential-move games with N players. In such settings, every player measures the kindness he infers from the actually chosen strategies of every player who played before him. The outcome of each of these individual comparisons can then be added up (or even scaled in a weighted average), in order to evaluate player i's distance function. Despite the greater generality of such model, nonetheless, its results and intuition are already captured by the two-player setting we consider in this paper.

4.4.1 Remarks on inequity aversion and reciprocity

In this paper we analyze how the consideration of foregone options affects players' equilibrium strategies. Nonetheless, in this subsection, we show that (under certain conditions) our model can also support the results of the literature on inequity aversion and intentions-based reciprocity as special cases.

Proposition 3. *Assume $s_j^{R_i}(s_i, s_j) = s_i$ for all s_j. Then, player i's preferences can be represented as a weighted average of her material payoffs and those of player j.*

$$U_i^C(s_i, s_j) = \gamma_i U_i^{NC}(s_i, s_j) + \gamma_j U_j^{NC}(s_j, s_i) \quad \text{where } \gamma_i, \gamma_j \in \mathbb{R} \tag{4.6}$$

In particular, the above proposition uses Segal and Sobel's (1999) results to specify that, when player i compares player j's actually chosen action, s_j, with that chosen by herself, s_i, her utility function $U_i^C(s_i, s_j)$ can be represented as an (additively separable) weighted average of both players' material payoffs. Therefore, in such context our model captures players' concerns for inequity aversion (or altruism) as a special case. In addition, this model also captures the literature on intentions-based reciprocity as a special case. Indeed, the above utility representation embodies Charness and Rabin's (2002) model for the case that player i infers misbehavior from player j's actions, and for $\gamma_i = 1 - \theta$ and $\gamma_j = -\theta$. That is,

$$\begin{aligned} U_i^C(s_i, s_j) &= (1-\theta)\, U_i^{NC}(s_i, s_j) - \theta U_j^{NC}(s_j, s_i) \\ &= U_i^{NC}(s_i, s_j) + \theta\left[U_i^{NC}(s_i, s_j) - U_j^{NC}(s_j, s_i)\right] \end{aligned} \tag{4.7}$$

Therefore, when players use their own action s_i as a reference point to compare other players' actually chosen action, s_j, our model embeds both inequity aversion and intentions-based reciprocity as special cases.[9]

[9] Clearly, this representation of player i's utility function does not *completely* capture Charness and Rabin's (2002) model, since they analyze other facets of individuals' behavior, such as inequity aversion, in addition to reciprocity. However, when restricted to intensions-based reciprocity alone, and when player i infers misbehavior from player j's actions, the above utility function coincides with that in Charness and Rabin (2002).

4.5 APPLICATIONS

4.5.1 Ultimatum bargaining game

Let us first apply our model to the ultimatum bargaining game where a (female) proposer j is called to choose how to divide a pie (of size normalized to one) between the (male) responder i and herself, and the responder either accepts or rejects the division suggested by the proposer, $s_i \in \{A, R\}$. In particular, let $(s_j, 1 - s_j)$ represent the actual division offered by player j, where s_j denotes the share of the pie accruing to the responder (which coincides with his payoff, $s_j = x_i$), and let $1 - s_j$ be the remaining share of the pie that the proposer keeps for herself (which coincides with the proposer's payoff, $1 - s_j = x_j$). Hence, x_i represents the offer that the proposer makes to the responder, and f_i denotes the foregone offer that the responder uses as a reference action, s_j^R. Specifically, the responder's utility function we use is given by the following expression[10], for any $x_i \in [0, 1]$, and $\alpha_i \geq 0$,

$$U_i^C(s_i, s_j) = s_j + \alpha_i \left(s_j - s_j^R\right) = x_i + \alpha_i (x_i - f_i) \tag{4.8}$$

Clearly, if $x_i > f_i$, the responder perceives kindness from the proposer, and gets his utility level increased in the second term. This additional utility is, furthermore, increasing in α_i, the parameter reflecting the importance that the responder assigns to the distance $x_i - f_i$. Intuitively, perceiving kind actions has greater effects on a receiver who is highly concerned about foregone options than on a receiver with small concerns about them. In addition, when either $\alpha_i = 0$ or $x_i = f_i$, the receiver's utility function just coincides with his utility when he is not concerned about the proposer's foregone options. In contrast, when $x_i < f_i$ the second term becomes negative. Now, the responder gets his utility level decreased from the unkindness he perceives from the proposer's actual offer, since $x_i < f_i$. Next, we check that the responder's utility function satisfies all the assumptions we consider in the previous section.

Lemma 3 $U_i^C(s_i, s_j)$ *satisfies A1 through A5.*

[10] Different functional forms for $U_i^C(s_i, s_j)$ satisfy assumptions A1 through A5, leading to the results predicted in the previous section. Nonetheless, a simple expression is used here to emphasize intuition.

We now introduce an example, in order to illustrate the main intuition behind the above utility function. In particular, we focus on the comparison between those utility functions analyzed in the literature and that suggested above, by using Brandts and Solà (2001) experimental results.

Example 1

Let us take an ultimatum bargaining game where the proposer chooses among two alternative divisions of the pie: $(x_i, x_j) = (0.2, 0.8)$ versus $(f_i, f_j) = (0.125, 0.875)$, where the aforementioned experimental results observe an overall accepting behavior from the receiver, or $(x_i, x_j) = (0.2, 0.8)$ versus $(f_i, f_j) = (0.875, 0.125)$, where the above experiments found several rejections. We first show that this pattern of rejections cannot be explained by Fehr and Schmidt's (1999) model on social preferences. When the receiver experiences inequity aversion, and the proposer offers $(x_i, x_j) = (0.2, 0.8)$ instead of $(f_i, f_j) = (0.125, 0.875)$, the receiver accepts if

$$x_i - \alpha_i \max\{x_j - x_i, 0\} - \beta_i \max\{x_i - x_j, 0\} =$$

$$0.2 - \alpha_i \max\{0.8 - 0.2, 0\} = 0.2 - 0.6\alpha_i > 0, \text{ if and only if } \alpha_i < \frac{1}{3}$$

While, in the case of receiving an offer $(x_i, x_j) = (0.2, 0.8)$ instead of $(f_i, f_j) = (0.875, 0.125)$ the receiver rejects if

$$0.2 - \alpha_i \max\{0.8 - 0.2, 0\} = 0.2 - 0.6\alpha_i < 0, \text{ if and only if } \alpha_i > \frac{1}{3}$$

which is not possible. Hence, this pattern of rejections cannot be explained by inequity aversion.

Let us now apply these payoffs to the utility function of the receiver with positive concerns about foregone options. In the case of receiving offer $(x_i, x_j) = (0.2, 0.8)$ instead of $(f_i, f_j) = (0.125, 0.875)$ the receiver accepts such offer if $0.2 + \alpha_i(0.2 - 0.125) = 0.2 - 0.075\alpha_i > 0$, i.e., $\alpha_i > -8/3$, which is satisfied since $\alpha_i \geq 0$. Similarly, applying it to the case in which the proposer offers $(x_i, x_j) = (0.2, 0.8)$ and foregoes $(f_i, f_j) = (0.875, 0.125)$, the receiver rejects it if $0.2 + \alpha_i(0.2 - 0.875) = 0.2 - 0.675\alpha_i < 0$, i.e., $\alpha_i > 0.29$. Thus, this offer is rejected if and only if the receiver's concern about foregone options is sufficiently high, $\alpha_i > 0.29$. ▲

Hence, the above utility function is then able to explain why an individualistic responder —who has no concerns about social payoffs— accepts an offer when it is associated to kindness from the proposer, $x_i > f_i$, but can reject this *same* offer when he evaluates it as a signal of unkindness. From the above utility function, we obtain the following result, describing the responder acceptance rule in this example.

Lemma 4. *In the ultimatum bargaining game with a responder who assigns a weight $\alpha_i \geq 0$ to the proposer's foregone divisions of the pie, f_i, the responder accepts any offer x_i if and only if $x_i \geqslant \bar{x}_i$, where $\bar{x}_i = \frac{\alpha_i}{1+\alpha_i} f_i$.*

Let us emphasize some interesting insights from the above lemma, illustrated in figure 6 below. Clearly, when $\alpha_i = 0$ the responder's acceptance rule collapses to $\bar{x}_i = 0$. Indeed, when the responder does not assign any weight to the proposer's unchosen actions, then any positive division of the pie is accepted by the responder, as in standard ultimatum bargaining games. Furthermore, the responder's acceptance threshold $\bar{x}_i$ is increasing in α_i, the importance he associates to the proposer's unchosen alternatives, i.e., he becomes more demanding in α_i. Finally, $\bar{x}_i$ is increasing in f_i, the receiver's foregone option (represented by an upward shift in the figure). Thus, the more demanding the receiver becomes (higher f_i) the more the proposer must offer him to induce his acceptance. Importantly, note that the minimum division that the receiver accepts, $\bar{x}_i$, is smaller than one (the total size of the pie) for any parameter values. Hence, $\bar{x}_i$ leaves some strictly positive portion of the pie to the proposer even when the receiver is extremely demanding (high α_i and f_i).

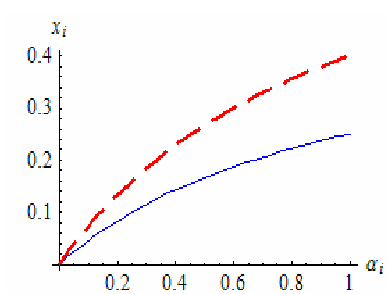

Figure 6: Lemma 4

Intuitively, the above acceptance rule of the responder shows that now the responder is *not* going to accept any positive offer, as the standard ultimatum bargaining game predicts when no concerns about the proposer's foregone options are considered. This fact clearly affects the proposer's optimal strategies. Certainly, if the proposer wants to obtain any positive payoff from the game, she must make an offer which is accepted by the responder, as we show below.

Proposition 4. *In the ultimatum bargaining game where the responder assigns an importance of* $\alpha_i \geq 0$ *to the options that the proposer forwent, the following strategy profile describes the unique subgame perfect equilibrium of the sequential game.*

Responder accepts any offer x_i *such that* $x_i \geqslant \bar{x}_i$, *where* $\bar{x}_i = \frac{\alpha_i}{1+\alpha_i} f_i$.

Proposer offers $x_i^* = \frac{\alpha_i}{1+\alpha_i} f_i$, *for any parameter values.*

Unlike models where the receiver is not concerned about foregone options —where the proposer keeps the entire pie for himself— the distribution of equilibrium payoffs when the receiver assigns a positive importance to foregone options is less unequal, as the following corollary specifies.

Corollary 1. *The distribution of equilibrium payoffs in the ultimatum bargaining game where the responder assigns importance* α_i *to the proposer's foregone option,* f_i, *is*

$$(x_i, x_j) = \left(\frac{\alpha_i}{1+\alpha_i} f_i, 1 - \frac{\alpha_i}{1+\alpha_i} f_i \right) \tag{4.9}$$

Indeed, note that this distribution of payoffs is more egalitarian than that of models where the receiver is not concerned about foregone options, $(x_i, x_j) = (0, 1)$, for any parameter values. Hence, by considering the proposer's foregone options into the responder's utility function we obtain higher degrees of fairness in the equilibrium payoffs, as well as higher cooperation between the players.

Let us finally relate our theoretical results with those of the experimental literature. In particular, Falk *et al.* (2003) and Brandts and Solà (2001) show the existence of a relationship between the receiver's acceptance threshold and the particular foregone offer that the proposer did not make. Indeed, both of these studies show that, conditional on offer $(x_i, x_j) = (0.2, 0.8)$ being made, the acceptance rate increases in the distance between the proposer's chosen and unchosen alternatives, as the following figures illustrate.

In particular, note that the first column of figure 7, where $x_i - f_i = 0.2 - 0.5 = -0.3$, represents a *negative* distance between the proposer's actual and foregone offer, from which the receiver infers "unkindness." On the other hand, column 3, where $x_i - f_i = 0.2 - 0 = 0.2$ (and the distance is *positive*) denotes the case in which the receiver interprets "kindness" from the proposer's offer, since she could have offered him less than she actually did. Finally, column 2 illustrates the case in which the proposer has no degree of freedom in choosing her particular offer to the receiver. i.e., the proposer's offer is (0.2,0.8) and her alternative is also (0.2,0.8). In this case, the outcome of the distance function is zero, what leads the receiver to neither perceive "kindness" nor "unkindness" from the proposer's actions.[11]

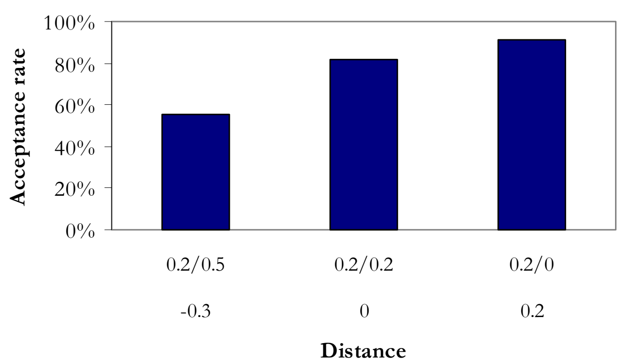

Figure 7: Falk et al. (2003)

Interestingly, the fact that the acceptance rate in the second column is exactly higher than when he perceives "unkindness" (column 1) but lower than when he infers "kindness"

[11] According to Falk *et al.* (2003), the small (but positive) percentage of rejections in this case can be supported by players' inequity aversion, since they might dislike the unequal payoff distribution resulting from their acceptance of (0.2,0.8). The fact that the responder does not attribute any responsibility to the proposer in settings where the latter does not have any choice to make has been extensively studied by psychologists with the use of attribution theory; see Ross and Fletcher (1985).

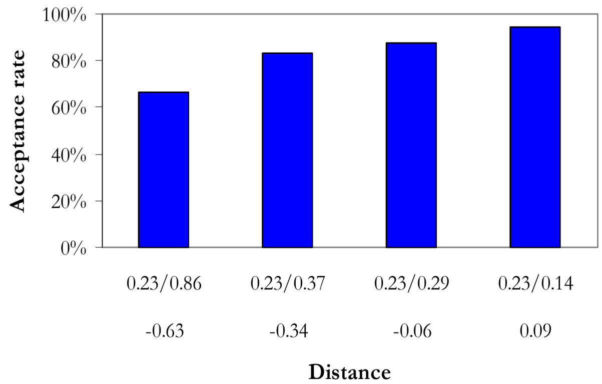

Figure 8: Brandts and Sola (2001)

(column 3) supports our results.[12] A similar intuition is also applicable to Brandts and Solà's (2001) results as figure 8 suggests. Hence, both of these studies confirm our theoretical prediction about the proposer's offer. Indeed, proposers are observed to make low offers when kindness can be inferred from such offers (positive distances), and high offers when they are interpreted in terms of unkindness (negative distances).

4.5.2 Labor market gift exchange game

We now apply the above model to a labor market gift exchange game, where the proposer is identified as a firm making a wage offer to a worker, who decides what level of effort to exert. In traditional models without considerations about unchosen options, since effort is costly and the worker is the last player to move, the worker's equilibrium strategy (in the subgame where the worker is called to move) is to exert zero effort regardless of the actual wage offer made by the firm. Operating by backwards induction, the subgame perfect equilibrium of this game predicts that the firm offers the lowest possible wage and that workers exert zero effort for any wage offered. These models, although theoretically simple, have found very limited experimental evidence. Indeed, Fehr and Gachter (2000) summarize a series

[12] Despite the regularity of their results (acceptance rates which increase in the outcome of the distance function), both of these studies report relatively high acceptance rates when distances are highly negative. Nonetheless, such acceptance rates are still lower than in the case of positive distances.

of experiments on labor markets where they confirm the existence of a positive correlation between the wage offered by the firm and the effort exerted by the worker.

We next suggest a utility function that satisfies the properties considered in section 3 and that can better rationalize the above experimental results. As in previous sections, we assume that the firm chooses a wage offer $x_i \in [0,1]$ to the worker. Similarly, let $f_i \in [0,1]$ represent the foregone wage offer that the worker uses (s_j^R) as a comparison against the actual wage offer x_i. In particular, let us consider the following utility function for the worker.

$$U_i^C(s_i, s_j) = s_j - e^2 + \alpha_i(s_j - s_j^{R_i})e = x_i - e^2 + \alpha_i(x_i - f_i)e \tag{4.10}$$

The above utility function coincides with the standard utility function of a worker who exerts costly effort when the parameter denoting the importance of foregone options, α_i, approaches zero. The third term represents the relevance of the foregone options for the worker, i.e., the wage offers that the firm did not make when proposing the actual offer x_i. Note that when the foregone wage proposal is higher than the actual wage offered, $x_i < f_i$, then this third term becomes negative, and the worker experiences a disutility from each unit of additional effort exerted. Similarly, when $x_i > f_i$, this third term becomes positive, and the worker interprets that the intentions of the firm are cooperative. That is, the worker observes that the firm offered a wage level which is above its foregone option, which in turn increases the worker's utility since he feels treated generously. In particular, this utility function for the worker satisfies all the assumptions we considered in section 3, as the following lemma specifies.

Lemma 5. *$U_i^C(s_i, s_j)$ satisfies A1 through A5.*

Intuitively, we should expect that, for proposals with a foregone option below the actual offer, the worker should feel pleased by the kindness of the firm, and responds by exerting a positive level of effort, in contrast to the standard game-theoretic model. These intuitions are confirmed in the following lemma.

Lemma 6. *In the gift exchange game where the worker assigns a value α_i to the distance between the firm's actual wage offer and its forgone alternative, the worker's optimal effort level (in the subgame induced after the wage proposal) is given by*

$$e\left(x_i\right) = \max\left\{\frac{1}{2}\alpha_i\left(x_i - f_i\right), 0\right\} \tag{4.11}$$

This optimal effort level is then positive if and only if the wage offer x_i is above the comparative foregone option, $x_i > f_i$, for any positive weight to foregone options, α_i. In addition, an increase in the relative importance that the worker assigns to foregone options increases his optimal effort level, i.e., $e\left(x_i\right)$ weakly increases in α_i. On the other hand, for a given weight on foregone options, α_i, and for a given wage offer x_i, optimal effort $e\left(x_i\right)$ increases as the comparative foregone option f_i decreases. Indeed, if the worker compares the actual wage he receives, x_i, with respect to the worst wage offer that the firm manager could ever pay him (e.g., the legal minimum wage), he is easily pleased by many positive wage offers. On the contrary, a worker who compares his relative position with respect to the best wage offer that the firm could afford to pay him certainly evaluates most of the wage offers he receives as a signal of unkindness from the firm manager.

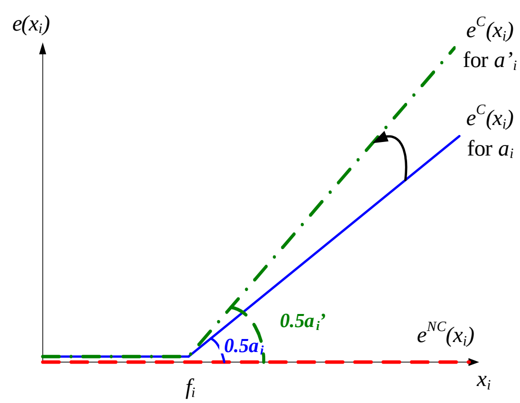

Figure 9: An increase in alpha i

This optimal effort level is illustrated in figures 9 and 10, which include in addition, the worker's effort level $e^{NC}\left(x_i\right)$ in the case of assigning no importance to foregone options. Note

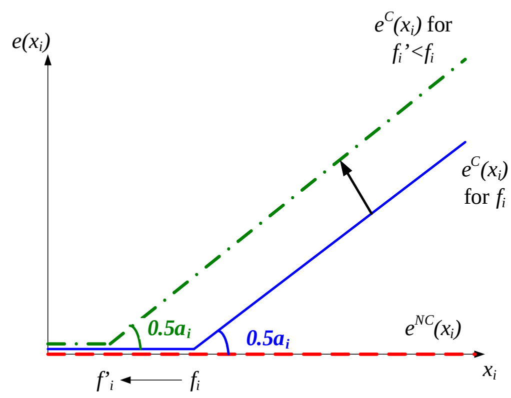

Figure 10: A decrease in f_i.

that $e^{NC}(x_i)$ is flat at zero for all x_i, since the worker exerts no effort for all wage offers. In both figures, the worker concerned about foregone options exerts positive efforts as long as $x_i > f_i$ for any positive weight on foregone options.[13] On the one hand, figure 9 indicates how the worker effort pivots upward —with center at $x_i = f_i$— when his concerns α_i about the firms' unchosen alternatives increase. On the other hand, figure 10 represents how the worker effort shifts upwards when the firm's unchosen alternative decreases (leftward shift in the horizontal intercept).

Interestingly, these results are not only supported by the aforementioned experimental evidence, but also by recent empirical work. In particular, Mas (2006) shows that police arrest rates and average sentence length decline (and crime reports raise) when the wage increase that police unions obtain is lower than their wage demands, relative to when it is higher. Hence, police union wage demands would work as the reference point which they use in their negotiations for higher salaries with government officials.

Given the above optimal effort function, and operating by backwards induction, we can find the firm's optimal wage offer. Specifically, we assume the following (standard) utility

[13] Note that our results in the labor market gift exchange game are similar to those in Akerlof (1982) since higher salaries induce higher effort levels. In particular, Akerlof's (1982) results are a special case of ours. when the foregone wage offer is exactly fixed at the "fair wage" level.

function for the firm, $V(s_j, s_i) = (v - x_i)\, e$, where v represents the constant productivity of effort (e.g., how worker's effort is transformed into final output); and x_i denotes, as above, the actual wage offer made to the worker. Moreover, $v > 1$, since the productivity of effort is assumed to be higher than any of the wage offers, $x_i \in [0, 1]$. Inserting the worker's optimal effort function found above, and manipulating, we find the optimal offer made by the firm.

Proposition 5. *In the gift exchange game where the worker assigns an importance of α_i to the distance between the wage offer foregone by the firm and its actual offer, the subgame perfect equilibrium strategies are the following*

Firm offers

$$x_i^* = \frac{v + f_i(x_i^*)}{2} \tag{4.12}$$

Worker accepts any offer x_i such that $x_i > 0$. In addition, the worker exerts an effort level of

$$e(x_i) = \max\left\{\frac{1}{2}\alpha_i (x_i - f_i(x_i)), 0\right\} \tag{4.13}$$

As the above proposition specifies, the firm's optimal offer x_i^* is higher than the worker's foregone option, $f_i(x_i^*)$, since $v > 1$. In addition, x_i^* is increasing in the foregone option, $f_i(x_i^*)$, that the receiver uses to make the comparison with respect[14] to x_i^*. In the standard models where concerns for foregone options are not considered, the subgame perfect equilibrium of the game predicts that the worker exerts no positive effort for any wage offer, and the firm, anticipating the worker's move, offers the lowest possible wage. In contrast, in the above environment including the importance of the foregone wage offers for the worker, we found that the firm makes a positive wage offer, since this offer can induce a higher level of exerted effort from the worker. That is, by showing kindness in high wage offers, the firm pleases the worker enough to induce him to exert higher efforts.

Clearly, the above equilibrium predictions are closer to the actual experimental results observed in the literature, Fehr and Gachter (2000), which find a positive correlation between

[14] Note that, for simplicity, we assume that the worker compares all wage offers with respect to the same foregone option, i.e., $f'(x_i^*) = 0$. Similar results are nonetheless applicable for the more general case in which $f'(x_i^*) \neq 0$, and they are included in the proof of proposition 5 at the appendix.

the wage offered by the firm and the exerted effort levels from the worker. Many authors have rationalized the above findings by using the *efficiency wage theory* arguments. That is, if a worker is paid above the minimum wage, he has a greater opportunity cost of shirking, which induces him to work harder, and to exert effort levels that are increasing in his wage offer. This paper may thus complement this rationalization of the experimental results through efficiency wage theory. Nonetheless, the model we presented above can explain cooperative behavior between employers and workers in the labor market without relying on the worker's opportunity cost of shirking, or his outside options if he is fired.

Finally, these results also provide an interesting explanation for the existence of *wage differentials* across industries. Indeed, as Krueger and Summers (1988) show, industry wage differentials are significant even after controlling for individual characteristics and firm quality; which suggests that these differentials are not just due to unobserved differences in labor quality. Our model then rationalizes this result by predicting that firms' equilibrium wage offer, after controlling for worker's productivity, may vary depending on the particular reference point that each worker uses for comparison.

4.5.3 Sequential public good game

The third game where we introduce the importance of the proposer's foregone options is the sequential public good game (PGG thereafter). Specifically, we consider a sequential solicitation game where a first mover is asked to submit a donation, $s_j \in [0,1]$, for the provision of a public good, and observing her donation, a follower decides which is the contribution, $s_i \in [0,1]$, he makes. In order to be consistent with the games defined above, the leader is assumed to not assign any weight to the follower's unchosen actions. In contrast, the follower assigns a relevance α_i to a specific contribution that the leader forwent, and that the follower uses as a reference point for comparison (reference action, s_j^R). In particular, leader and follower's utility functions are, respectively

$$U_j^{NC}(s_j, s_i) = z_j + [m(s_i + s_j)]^{0.5}$$

$$U_i^{C}(s_i, s_j) = z_i + \left[m(s_i + s_j)\left[1 + \alpha_i\left(s_j - s_j^R\right)\right]\right]^{0.5} \tag{4.14}$$

Both of these functions are quasilinear in the private good, z, and their nonlinear part takes into account the utility derived from the total public good provision $G = s_i + s_j$ (relevant for both players) and the distance $\alpha_i\left(s_j - s_j^R\right)$, which is only relevant for the follower. For simplicity, let us assume in this application that the follower uses the same reference action s_j^R for all action choices of the leader. Finally, $m \geq 0$ denotes the return every player obtains from total contributions to the public good. Interestingly, note how foregone options are introduced into the follower's utility function. When the relevance he assigns to the leader's unchosen alternatives approaches zero, $\alpha_i = 0$, the follower only cares about the private and public good consumption. However, when he assigns a positive importance to foregone options, he experiences a higher utility from contributing to the public good when the leader's contribution is higher than the foregone option, $s_j > s_j^R$, or a lower utility otherwise, $s_j < s_j^R$. In addition, this utility function satisfies all the assumptions we consider in section 3, as the following lemma states.

Lemma 7. *$U_i^C(s_i, s_j)$ satisfies A1 through A5.*

Since we are discussing a sequential game where the follower decides how much to give out of a continuous strategy choice, the second mover best response function is easily found by solving the follower's utility maximization problem. We summarize this result in the following lemma.

Lemma 8. *In the sequential PGG, where the follower assigns weight α_i to the distance between the leader's actual contribution, s_j, and the foregone contribution, s_j^R, the follower's best response function $s_i^C(s_j)$ is given by*

$$s_i^C(s_j) = \begin{cases} \frac{m\left(1-\alpha_i s_j^R\right)}{4} - \left(1 + \frac{\alpha_i m}{4}\right) s_j \ \ if \ \ s_j \in \left[0, \frac{m\left(1-\alpha_i s_j^R\right)}{4-\alpha_i m}\right) \\ 0 \ \ if \ \ s_j \geq \frac{m\left(1-\alpha_i s_j^R\right)}{4-\alpha_i m} \end{cases} \tag{4.15}$$

Figure 11 compares the second mover's best response function when he is concerned about foregone options, $s_i^C(s_j)$, and when he is not, $s_i^{NC}(s_j)$.

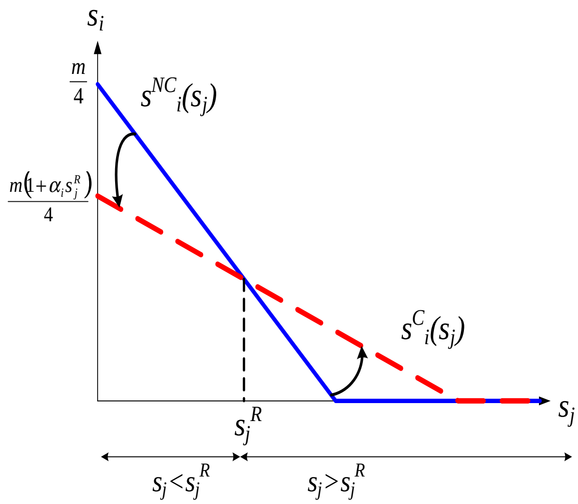

Figure 11: Comparing $s_i^C(s_j)$ and $s_i^{NC}(s_j)$

Specifically, note that the introduction of the importance of foregone options into the second mover's utility function makes $s_i^C(s_j)$ to pivot counterclockwise with respect to $s_i^{NC}(s_j)$, with center at $s_j = s_j^R$, making $s_i^C(s_j)$ steeper than $s_i^{NC}(s_j)$. Hence, the second mover relatively "reciprocates" the first mover's contributions, since he reduces his donation when $s_j < s_j^R$, but increases it when $s_j > s_j^R$. After finding $s_i^C(s_j)$, and by sequential rationality, we can now find the first mover's equilibrium contribution in this game.

Lemma 9. *In the sequential PGG, where the follower assigns a weight α_i to the leader's foregone options, the leader's donation in the subgame perfect Nash equilibrium of the game is*

$$s_j^* = \begin{cases} 0 \quad \textit{if} \quad \alpha_i < \bar{\alpha}_i \\ \frac{16\left(\alpha_i s_j^R - 1\right) + \alpha_i^2 m^2}{16\alpha_i} \quad \textit{otherwise} \end{cases} \tag{4.16}$$

where $\bar{\alpha}_i = \frac{16}{16 s_j^R + m}$

Thus, the first donor submits a zero contribution when the second donor's concerns for foregone options are low enough, $\alpha_i < \bar{\alpha}_i$. Clearly, when $\alpha_i = 0$ the first donor also submits a null donation, which coincides with the equilibrium prediction in standard PGGs. However, when the second donor's concerns for foregone options increase enough, $\alpha_i > \bar{\alpha}_i$, the first mover is induced to submit positive contributions that can trigger further donations from the second mover (given his reciprocating behavior described in the previous figure). Additionally, note that as expected, the leader's contribution is increasing in the follower's concerns for foregone options, α_i, and in the foregone contribution that he uses as a reference point for comparison, s_j^R.

Proposition 6. *In the sequential PGG where the second mover assigns a weight α_i to the first mover's unchosen alternatives, the following strategy profile describes the subgame perfect equilibrium of the game.*

Proposer contributes

$$s_j^* = \begin{cases} 0 \;\; \text{if} \;\; \alpha_i < \bar{\alpha}_i \\ \frac{16\left(\alpha_i s_j^R - 1\right) + \alpha_i^2 m^2}{16\alpha_i} \;\; \text{otherwise} \end{cases} \tag{4.17}$$

And *the second mover responds by contributing*

$$s_i^C(s_j) = \begin{cases} \frac{m\left(1-\alpha_i s_j^R\right)}{4} - \left(1 + \frac{\alpha_i m}{4}\right) s_j \;\; \text{if} \;\; s_j \in \left[0, \frac{m\left(1-\alpha_i s_j^R\right)}{4-\alpha_i m}\right) \\ 0 \;\; \text{if} \;\; s_j \geq \frac{m\left(1-\alpha_i s_j^R\right)}{4-\alpha_i m} \end{cases} \tag{4.18}$$

Particularly, the above results specify that by having a second mover concerned about the first mover's foregone options, the latter is induced to contribute (weakly) higher amounts than those she would donate in the case of facing a responder with no concerns about her unchosen alternatives. From a more general perspective, by introducing a follower concerned about the leader's foregone options, we are able to obtain (weakly) higher levels of cooperation in the public good provision.

4.6 CONCLUSIONS

Different experimental papers, such as Brandts and Solà (2001), Falk *et al.* (2003), and Andreoni *et al.* (2002), accumulate a significant evidence about the importance of a player's unchosen alternatives on other players' actions. Foregone options, in particular, may work as standards against which every individual evaluates the kindness of other players in the population. Importantly, these studies suggest that arguments on social preferences alone cannot explain their experimental results without complementing their approach by considering the importance of a players' unchosen alternatives inside his opponents' utility function.

This paper examines a tractable theoretical model that introduces these unchosen alternatives into individuals' preferences via a reference point. We first analyze the equilibrium prediction in complete information sequential-move games, and then compare it with that of standard games where players are not concerned about unchosen alternatives. We show that, without relying on interpersonal payoff comparisons (i.e., within "strictly individualistic" preferences), our model predicts higher levels of fairness in the resulting allocation, as well as higher cooperation among the players, than standard game-theoretic models. In addition, we demonstrate that this approach embeds as special cases many existing behavioral models: from inequity aversion to intentions-based reciprocity. Therefore, this model offers a broader and more unifying explanation of agents' conduct than these models alone. Furthermore, when applying our model to different sequential games, we obtain interesting results. First, the equilibrium allocation in the ultimatum bargaining game is fairer than that resulting from standard game-theoretic predictions. Second, worker's effort and firm's proposed wages are higher than in the usual labor market gift exchange model. Finally, equilibrium donations in the sequential public good game are higher than the predictions for standard models.

There are several natural extensions to the model introduced in this paper. First, it would be interesting to experimentally test under which payoff structures we can rationalize observed behavior using individuals' preferences over equitable payoffs, and in which environments human conduct is instead mainly explained by the players' "strictly individualistic preferences" suggested in this paper. One direct test of the dominance of these two behavioral motives is, for example, the following ultimatum bargaining game. The proposer is allowed to make only two divisions of the pie, of size normalized to one. In the first treatment she can offer (0.4, 0.6), giving 0.4 to the responder and keeping 0.6 for herself, or the equitable payoff (0.5, 0.5). In the second treatment, the first division of the pie is fixed in (0.4, 0.6), but the second division is now (0.6, 0.4) instead. Note that, conditional on the first offer, (0.4, 0.6), being made, the distance between the actual offer, 0.4, and the alternative offer is higher in the first treatment, $0.4 - 0.5 = -0.1$, than in the second, $0.4 - 0.6 = -0.2$. Hence, according to our equilibrium predictions, we should observe more rejections in the second treatment than in the first. However, if we observe higher percentage of rejections in the first than in the second treatment, it must be that responders in the first treatment evaluate the equitable payoffs that the proposer did not select as a more desirable goal than the higher individual payoff he could have received in the second treatment.

Second, in this paper the space of available alternatives was exogenously determined before the beginning of the game. However, it would be interesting to allow players to strategically select their available choices before the game starts, given that the kindness other players perceive from their chosen actions depends on which available strategies are *not* chosen. That is, by strategically selecting her set of available alternatives, a player may achieve that other players infer a greater kindness from her actions. This strategic selection of available choices is observed in different contexts, where a player uses one of her unchosen alternatives as an excuse to support her actual choices, since the equilibrium payoff associated with that particular unchosen action would have been certainly worse than that from her chosen action. These extensions can certainly enhance our understanding of the role of players' foregone options on their opponents' incentives, and how such incentives can lead to higher degrees of cooperation from a strictly individualistic perspective.

4.6.1 *Proof of Lemma 1*

In this environmental game, both players are asked to simultaneously submit their investments in emission-reducing technologies. Fixing country j's investment, x_j, country i's utility maximization problem becomes

$$\max_{x_i} \quad w - x_i + \ln\left[m(x_i + x_j) + \alpha_i\left(x_i - c_i\right)\right]$$

And the argument that maximizes this utility function gives us the following best response function

$$x_i(x_j) = \begin{cases} \frac{1}{m+\alpha_i}\alpha_i c_j \text{ if } x_j = 0 \\ 1 + \frac{1}{m+\alpha_i}\left[\alpha_i c_i - m x_j\right] \text{ if } x_j \in \left]0, \frac{\alpha_i(1+c_i)+m}{m}\right[\\ 0 \text{ if } x_j \geqslant \frac{\alpha_i(1+c_i)+m}{m} \end{cases}$$

Since $1+\frac{1}{m+\alpha_i}\left[\alpha_i c_i - m x_j\right] = 0$ exactly at $x_j = \frac{\alpha_i(1+c_i)+m}{m}$. Hence, this best response function can be more compactly expressed as

$$x_i(x_j) = \begin{cases} 1 + \frac{1}{m+\alpha_i}\left[\alpha_i c_i - m x_j\right] \text{ if } x_j \in \left[0, \frac{\alpha_i(1+c_i)+m}{m}\right[\\ 0 \text{ if } x_j \geqslant \frac{\alpha_i(1+c_i)+m}{m} \end{cases}$$

4.6.2 *Proof of Proposition 1*

Let us take country i's best response function, $x_i(x_j)$, from lemma 1, and analyze the different forms in which country i and j's best response functions can cross each other. The corner solutions (cases 1 and 2) are illustrated in figures 12 and 13, to clarify the following steps of the proof.

Case 1: $x_i^* = 0$

Note that $x_i^* = 0$ if and only the following two conditions are satisfied: (1) the horizontal intercept of country i's best response function is lower than that of country j, and (2) the slope of country j's best response function is small enough to make that $x_j(x_i)$ does not cross $x_i(x_j)$. That is, the first condition is satisfied if

$$\frac{\alpha_i(1+c_i)}{m} + 1 < \frac{\alpha_j c_j}{\alpha_j + m} + 1$$

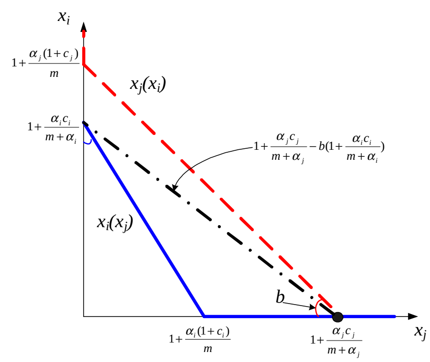

Figure 12: Proposition 1. Case 1

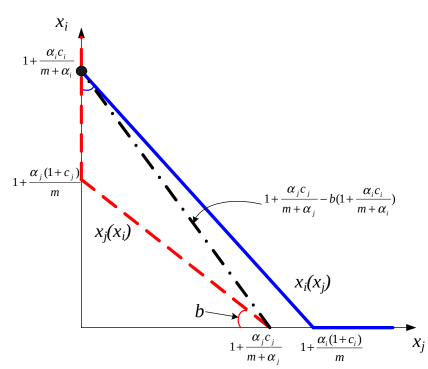

Figure 13: Proposition 1. Case 2

Manipulating this inequality, we obtain

$$\alpha_i < \frac{\alpha_j c_j m}{(1+c_i)(\alpha_j+m)}$$

On the other hand, the second condition holds if, b, the slope of country j's best response function, satisfies

$$0 < 1 + \frac{\alpha_j c_j}{m+\alpha_j} - b(1+\frac{\alpha_i c_i}{m+\alpha_i})$$
$$\iff b < \frac{[m+\alpha_j(1+c_j)][m+\alpha_i]}{[m+\alpha_i(1+c_i)][m+\alpha_j]}$$

and since the slope of $x_j(x_i)$ is $\frac{m}{\alpha_j+m}$, we need that

$$\begin{aligned} \frac{m}{m+\alpha_j} &< \frac{[m+\alpha_j(1+c_j)][m+\alpha_i]}{[m+\alpha_i(1+c_i)][m+\alpha_j]} \\ [m+\alpha_i(1+c_i)][m+\alpha_j]m &< [m+\alpha_j(1+c_j)][m+\alpha_i][m+\alpha_j] \end{aligned}$$

and manipulating, and solving for α_i, we obtain the threshold of α_i below which all values of α_i support a zero investment in clean technologies by country i,

$$\alpha_i \leq \frac{mc_j + \alpha_j(1+c_j)(m+c_j)}{(1+c_i)m}$$

Case 2: $x_i^* = 1 + \frac{\alpha_i c_i}{m+\alpha_i}$

Let us now analyze the case in which country i sets the maximum investment $(1+\frac{\alpha_i c_i}{m+\alpha_i})$, while country j does not invest. Firstly, we need that country i's horizontal intercept is above that of country j's, what simply implies

$$\begin{aligned} \frac{\alpha_i(1+c_i)}{m}+1 &> \frac{\alpha_j c_j}{\alpha_j+m}+1 \iff \alpha_i > \frac{\alpha_j c_j m}{(1+c_i)(\alpha_j+m)} \\ &\iff \alpha_i > \frac{\alpha_j c_j m}{(1+c_i)(\alpha_j+m)} = \bar{\alpha}_i(\alpha_j) \end{aligned}$$

Secondly, we need that b, the slope of country j's best response function, satisfies

$$0 > 1 + \frac{\alpha_j c_j + m}{m+\alpha_j} - b(1+\frac{\alpha_i c_i}{m+\alpha_i})$$

and operating similarly as in the previous case, we have

$$\alpha_i > \frac{mc_j + \alpha_j(1+c_j)(m+c_j)}{(1+c_i)m} = \hat{\alpha}_i(\alpha_j)$$

Case 3: $x_i^* = \frac{\alpha_i(1+c_i)(\alpha_j+m)-\alpha_j m c_j}{\alpha_j m+\alpha_i(\alpha_j+m)}$

Finally, the equilibrium is interior when first, country i's horizontal intercept is below that of country j's, what simply implies,

$$\frac{\alpha_i(1+c_i)}{m}+1<\frac{\alpha_j c_j}{\alpha_j+m}+1 \Longleftrightarrow \alpha_i<\frac{\alpha_j c_j m}{(1+c_i)(\alpha_j+m)}=\bar{\alpha}_i(\alpha_j)$$

and second, when b, the slope of country j's best response function, satisfies

$$0>1+\frac{\alpha_j c_j+m}{m+\alpha_j}-b(1+\frac{\alpha_i c_i}{m+\alpha_i}) \Longleftrightarrow \alpha_i>\frac{mc_j+\alpha_j(1+c_j)(m+c_j)}{(1+c_i)m}=\hat{\alpha}_i(\alpha_j)$$

Finally, we must check that $\bar{\alpha}_i(\alpha_j)>\hat{\alpha}_i(\alpha_j)$. Indeed,

$$\begin{aligned}\frac{\alpha_j c_j m}{(1+c_i)(\alpha_j+m)}-\frac{mc_j+\alpha_j(1+c_j)(m+c_j)}{(1+c_i)m} &> 0\\ \alpha_j &> -\frac{m}{(1+c_j)}\end{aligned}$$

Since m and c_j are positive this inequality holdsfor any parameter values. Hence, we can summarize the above three cases as follows:

$$x_i^*=\begin{cases} 1+\frac{\alpha_i c_i}{m+\alpha_i} \text{ if } \alpha_i>\bar{\alpha}_i(\alpha_j)\\ \frac{\alpha_i(1+c_i)(\alpha_j+m)-\alpha_j m c_j}{\alpha_j m+\alpha_i(\alpha_j+m)} \text{ if } \alpha_i\in]\hat{\alpha}_i(\alpha_j),\bar{\alpha}_i(\alpha_j)]\\ 0 \text{ if } \alpha_i\in]0,\hat{\alpha}_i(\alpha_j)]\end{cases}$$

where $\bar{\alpha}_i(\alpha_j)=\frac{\alpha_j c_j m}{(1+c_i)(\alpha_j+m)}$ and $\hat{\alpha}_i(\alpha_j)=\frac{mc_j+\alpha_j(1+c_j)(m+c_j)}{(1+c_i)m}$

4.6.3 *Proof of Lemma 2*

In order to obtain a higher investment level than in the case of unconcerned countries, we need to show that the total sum of the optimal investment in clean technologies of country i and j is greater than one. Notice that there are three possible combinations, which depend of the parameter α_i or α_j.

	x_i	x_j
Case 1	$\alpha_i \in]0, \hat{\alpha}_i(\alpha_j)]$	$\alpha_j > \bar{\alpha}_j(\alpha_i)$
Case 2	$\alpha_i \in]\hat{\alpha}_i(\alpha_j), \bar{\alpha}_i(\alpha_j)]$	$\alpha_j \in]\hat{\alpha}_j(\alpha_i), \bar{\alpha}_j(\alpha_i)]$
Case 3	$\alpha_i > \bar{\alpha}_i(\alpha_j)$	$\alpha_j \in]0, \hat{\alpha}_j(\alpha_i)]$

Case 1 and 3 are trivial because when $\alpha_i > \bar{\alpha}_i(\alpha_j)$ in case 1 (and $\alpha_j > \bar{\alpha}_j(\alpha_i)$ in case 3) the optimal investment is $x_i^* = 1 + \frac{\alpha_i c_i}{m+\alpha_i}$ in case 1 and $x_j^* = 0$ $(x_j^* = 1 + \frac{\alpha_j c_j}{m+\alpha_j}$ and $x_i^* = 0$ in case3) which clearly is greater than one.

In case 2 the optimal investments in clean technologies are $x_i^* = \frac{\alpha_i(1+c_i)(\alpha_j+m)-\alpha_j m c_j}{\alpha_j m+\alpha_i(\alpha_j+m)}$ and $x_j^* = \frac{\alpha_j(1+c_j)(\alpha_i+m)-\alpha_i m c_i}{\alpha_i m+\alpha_j(\alpha_i+m)}$, therefore:

$$
\begin{aligned}
x_i^* + x_j^* &> 1 \\
\frac{\alpha_i(1+c_i)(\alpha_j+m)-\alpha_j m c_j}{\alpha_j m+\alpha_i(\alpha_j+m)} + \frac{\alpha_j(1+c_j)(\alpha_i+m)-\alpha_i m c_i}{\alpha_i m+\alpha_j(\alpha_i+m)} &> 1 \\
\frac{\alpha_j m+\alpha_i(m+\alpha_j(2+c_i+c_j))}{\alpha_j m+\alpha_i(\alpha_j+m)} &> 1 \\
c_i + c_j &> -1
\end{aligned}
$$

and given that c_i and c_j are positive numbers the above inequality holds.

Therefore the sum of the optimal investments in case 2 is greater than one

4.6.4 *Proof of Lemma 3*

Differentiating x_i^* with respect to c_i,

$$\frac{\partial x_i^*}{\partial c_i} = \begin{cases} \frac{\alpha_i}{m+\alpha_i} \text{ if } \alpha_i > \bar{\alpha}_i(\alpha_j) \\ \frac{\alpha_i(\alpha_j+m)}{\alpha_j m+\alpha_i(\alpha_j+m)} \text{ if } \alpha_i \in]\hat{\alpha}_i(\alpha_j), \bar{\alpha}_i(\alpha_j)] \\ 0 \text{ if } \alpha_i \in]0, \hat{\alpha}_i(\alpha_j)] \end{cases}$$

which is positive for any parameter values. Similarly, differentiating x_i^* with respect to c_j,

$$\frac{\partial x_i^*}{\partial c_j} = \begin{cases} 0 \text{ if } \alpha_i > \bar{\alpha}_i(\alpha_j) \\ -\frac{\alpha_j m}{\alpha_j m+\alpha_i(\alpha_j+m)} \text{ if } \alpha_i \in]\hat{\alpha}_i(\alpha_j), \bar{\alpha}_i(\alpha_j)] \\ 0 \text{ if } \alpha_i \in]0, \hat{\alpha}_i(\alpha_j)] \end{cases}$$

which is negative for any parameter values.

4.6.5 *Proof of Lemma 4*

Differentiating x_i^* with respect to α_i,

$$\frac{\partial x_i^*}{\partial \alpha_i} = \begin{cases} \frac{c_i m}{[\alpha_i+m]^2} \text{ if } \alpha_i > \bar{\alpha}_i(\alpha_j) \\ \frac{\alpha_j m(c_i+c_j+1)(\alpha_j+m)}{[\alpha_j m+\alpha_i(\alpha_j+m)]^2} \text{ if } \alpha_i \in]\hat{\alpha}_i(\alpha_j), \bar{\alpha}_i(\alpha_j)] \\ 0 \text{ if } \alpha_i \in]0, \hat{\alpha}_i(\alpha_j)] \end{cases}$$

which is positive. Finally, let us differentiate x_i^* with respect to α_j,

$$\frac{\partial x_i^*}{\partial \alpha_j} = \begin{cases} 0 \text{ if } \alpha_i > \bar{\alpha}_i(\alpha_j) \\ -\frac{\alpha_i m^2(c_i+c_j+1)}{[\alpha_j m+\alpha_i(\alpha_j+m)]^2} \text{ if } \alpha_i \in]\hat{\alpha}_i(\alpha_j), \bar{\alpha}_i(\alpha_j)] \\ 0 \text{ if } \alpha_i \in]0, \hat{\alpha}_i(\alpha_j)] \end{cases}$$

which is negative.

4.6.6 *Proof of Proposition 2*

Inserting the results from proposition 1 for two countries with positive weight, $\alpha_i > 0$ and $\alpha_j > 0$, we obtain country i's equilibrium utility level from playing the subgame in which countries invest in emission-reducing technologies.

$$\begin{aligned} U_i &= w - x_i^* + \ln\left[m\left(x_i^* + x_j^*\right) + \alpha_i\left(x_i^* - c_i\right)\right] \\ &= -\frac{\alpha_i(1+c_i)(\alpha_j+m)+\alpha_j c_j m}{\alpha_j m + \alpha_i\left(\alpha_j+m\right)} + w - \ln\left[\alpha_i + m\right] \end{aligned}$$

Since U_i is linear in both c_i and c_j, we can determine the value of c_i and c_j that maximizes U_i by checking if U_i increases or decreases in c_i and c_j. Indeed,

$$\frac{\partial U_i}{\partial c_i} = -\frac{\alpha_i(\alpha_j+m)}{\alpha_j m + \alpha_i\left(\alpha_j+m\right)}$$

which is negative for all parameter values. In contrast,

$$\frac{\partial U_i}{\partial c_j} = \frac{\alpha_j m}{\alpha_j m + \alpha_i\left(\alpha_j+m\right)}$$

which is positive for all parameter values. Hence, U_i decreases in c_i and increases in c_j.

4.6.7 *Proof of Proposition 3*

In the ultimatum bargaining game country i makes an offer and country j can accept or reject such offer. In the model there are 3 cases, since countries i and j can have different concern levels about green voters. Specifically, country i solves the following maximization problem:

$$\begin{gathered} \underset{c_i,c_j}{Max} U_i(c_i,c_j) \\ s.t. U_j(c_i,c_j) = 0 \end{gathered}$$

First case: Country i's concern level is very high, $\alpha_i > \bar{\alpha}_i(\alpha_j)$, and country j's concern level is small, $\alpha_j \in]0, \hat{\alpha}_j(\alpha_i)]$.

$$
\begin{aligned}
(c_i^*, c_j^*) &= \left(\frac{\alpha_i(-1+m+\alpha_i)}{1+\alpha_i}, \frac{e^w(\alpha_i^3 e^w m + mB + \alpha_i(1+m)B + \alpha_i^2(-1+e^w m^2))}{\alpha_j(1+\alpha_i)(m+\alpha_i)}\right) \\
\text{where } B &= (-1+e^w m) \\
(x_i^*, x_j^*) &= \left(\frac{m+\alpha_i(1+m+\alpha_i(m+\alpha_i))}{(1+\alpha_i)(m+\alpha_i)}, 0\right)
\end{aligned}
$$

Second case: Country i and j have a medium concern level, that is $\alpha_i \in]\hat{\alpha}_i(\alpha_j), \bar{\alpha}_i(\alpha_j)]$ and $\alpha_j \in]\hat{\alpha}_j(\alpha_i), \bar{\alpha}_j(\alpha_i)]$.

$$
\begin{aligned}
c_i^* &\in \left[0, \frac{\alpha_i(-1+w)+m(-1+2w)+\Psi\log\Psi+m\log\Gamma)}{\alpha_i}\right] \text{ and} \\
c_j^* &\in \left[\frac{-\alpha_i\alpha_j+\alpha_j m(-1+w)+\alpha_i\Gamma w+(\alpha_j m+\alpha_i\Gamma)\log\Gamma}{\alpha_j\Psi},\right. \\
&\quad \left.\frac{\alpha_j(-1+w)+m(-1+2w)+\Gamma\log\Gamma+m\log\Psi)}{\alpha_j}\right] \\
\text{where } \Gamma &= (\alpha_j+m) \text{ and } \Psi = (\alpha_i+m) \\
x_i^* &\in \left[\frac{\Psi-mw-m\log\Gamma}{\Psi}, w+\log\Psi\right] \text{ and } x_j^* \in \left[\frac{\alpha_j(2\Psi-mw-m\log\Gamma)}{\alpha_j m+\alpha_i\Gamma}, w+\log\Gamma\right]
\end{aligned}
$$

Third case: Country i's concern level is very small, $\alpha_i \in]0, \hat{\alpha}_i(\alpha_j)]$, and country j is very concerned about the noncompliance cost, $\alpha_j > \bar{\alpha}_j(\alpha_i)$.

$$
\begin{aligned}
c_i^* &\in \left[0, \frac{e^{-w}(-1+e^w mw+e^w m\log[\alpha_j+m])}{\alpha_i}\right] \text{ and } c_j^* = \frac{(m+\alpha_j)(-1+w+\log[\alpha_j+m])}{\alpha_j} \\
x_i^* &= 0 \text{ and } x_j^* \in [1, w+\log[\alpha_j+m])
\end{aligned}
$$

Figure 14 depicts the relationship between countries' concern levels and the main results obtained in Proposition 3.

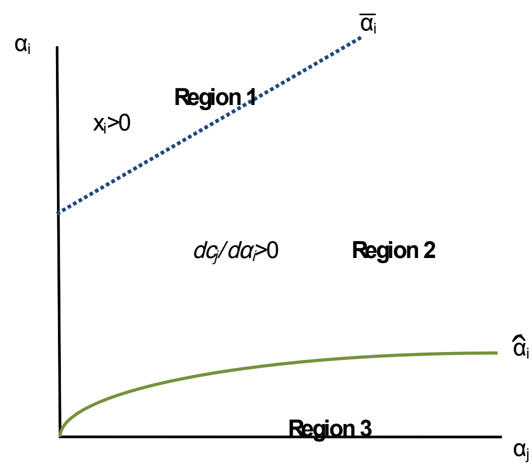

Figure 14: Results proposition 3

4.6.7.1 *Corollary 2* Case when country i proposes (c_i^*, c_j^*) and $\alpha_i > \bar{\alpha}_i(\alpha_j)$ and $\alpha_j \in]0, \hat{\alpha}_j(\alpha_i)]$:

$$
\begin{aligned}
(c_i^*, c_j^*) &= (0, \frac{e^w(-1+e^w m)}{\alpha_j}) \\
(x_i^*, x_j^*) &= (1, 0) \\
\text{where } c_i^* &= \frac{\alpha_i(-1+m+\alpha_i)}{1+\alpha_i} = 0 \text{ iff } \alpha_i \leq 1-m
\end{aligned}
$$

4.6.8 *Corollary 3*

Case when country i proposes (c_i^*, c_j^*) and $\alpha_i \in]\hat{\alpha}_i(\alpha_j), \bar{\alpha}_i(\alpha_j)]$ and $\alpha_j \in]\hat{\alpha}_j(\alpha_i), \bar{\alpha}_j(\alpha_i)]$:

$$\frac{\partial c_j}{\alpha_i} = \frac{m^2(w + \log \Gamma)}{\alpha_j \Psi^2} > 0$$

$$\frac{\partial c_j}{\alpha_i} = \frac{m}{\alpha_j \Psi} > 0$$

Case when country i proposes (c_i^*, c_j^*) and $\alpha_i > \bar{\alpha}_i(\alpha_j)$ and $\alpha_j \in]0, \hat{\alpha}_j(\alpha_i)]$:

$$\frac{\partial c_j}{\partial \alpha_i} = \frac{m\alpha_i(\alpha_i^3 + 2(-1+m)m + 2\alpha_i^2(1+m) + \alpha_i(-1+m(3+m)))}{\alpha_j(1+\alpha_i)^2(m+\alpha_i)^2} > 0, \text{ iff } m > 0$$

4.6.9 *Corollary 4*

The total investment in clean technology, $G = x_i + x_j$, obtained when countries' concern level are represented by $\alpha_i \in]\hat{\alpha}_i(\alpha_j), \bar{\alpha}_i(\alpha_j)]$ and $\alpha_j \in]\hat{\alpha}_j(\alpha_i), \bar{\alpha}_j(\alpha_i)]$ is:

$$\begin{aligned} G^* &= x_i^* + x_j^* = 2w + \log\Gamma + \log\Psi \\ \frac{\partial G^*}{\partial \alpha_j} &= \frac{1}{(\alpha_j + m)} > 0 \\ \frac{\partial G^*}{\partial \alpha_i} &= \frac{1}{(\alpha_i + m)} > 0 \end{aligned}$$

4.6.10 *Proof of Lemma 5*

In this environmental game, both players are asked to simultaneously submit their investments in emission-reducing technologies. Fixing country j's investment, x_j, country i's utility maximization problem becomes

$$\max_{x_i} \quad w - x_i + \ln\left[m(x_i + x_j) + \alpha_i\left(x_j - c_j\right)\right]$$

And the argument that maximizes this utility function gives us the following best response function

$$x_i(x_j) = \begin{cases} \frac{1}{m}\left[\alpha_i c_j + m\right] \text{ if } x_j = 0 \\ \frac{1}{m}\left[\alpha_i c_j + m - (\alpha_i + m)x_j\right] \text{ if } x_j \in \left]0, \frac{\alpha_i c_j + m}{\alpha_i + m}\right[\\ 0 \text{ if } x_j \geqslant \frac{\alpha_i c_j + m}{\alpha_i + m} \end{cases}$$

Since $\frac{1}{m}\left[\alpha_i c_j + m - (\alpha_i + m)x_j\right] = 0$ exactly at $x_j = \frac{\alpha_i c_j + m}{\alpha_i + m}$. Hence, this best response function can be more compactly expressed as

$$x_i(x_j) = \begin{cases} \frac{1}{m}\left[\alpha_i c_j + m - (\alpha_i + m)x_j\right] \text{ if } x_j \in \left[0, \frac{\alpha_i c_j + m}{\alpha_i + m}\right[\\ 0 \text{ if } x_j \geqslant \frac{\alpha_i c_j + m}{\alpha_i + m} \end{cases}$$

4.6.11 *Proof of Proposition 4*

Let us take country i's best response function, $x_i(x_j)$, from lemma 1, and analyze the different forms in which country i and j's best response functions can cross each other. The corner solutions (cases 1 and 2 below) are illustrated in figures 15 and 16, to clarify the following steps of the proof.

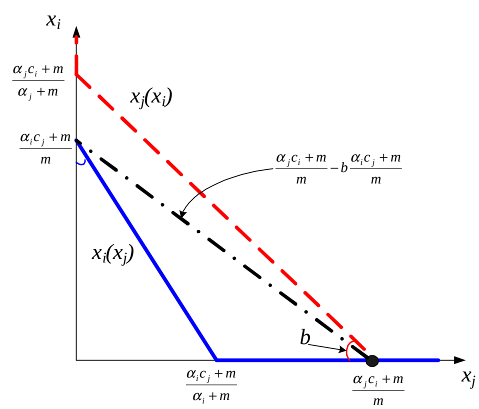

Figure 15: Proposition 4. Case 1

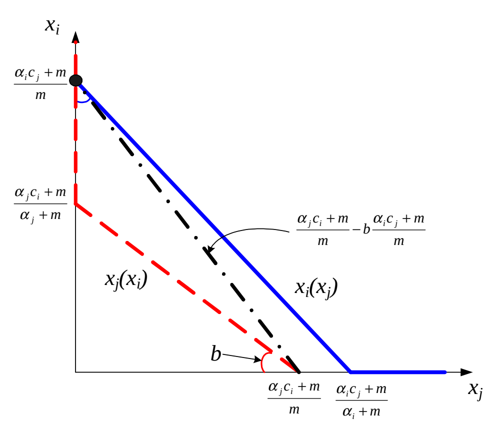

Figure 16: Proposition 4. Case 2

Case 1: $x_i^* = 0$

Note that $x_i^* = 0$ if and only the following two conditions are satisfied: (1) the horizontal intercept of country i's best response function is lower than that of country j., and (2) the slope of country j's best response function is small enough to make that $x_j(x_i)$ does not cross $x_i(x_j)$. That is, the first condition is satisfied if

$$\frac{\alpha_i c_j + m}{\alpha_i + m} < \frac{\alpha_j c_i + m}{m}$$

Manipulating this inequality, we obtain

$$\alpha_i < \frac{\alpha_j c_i m}{m\,(c_j - 1) - \alpha_j c_i}$$

and since $c_i \leq 1$ by definition, the term in the right-hand side is negative for any $c_i < 1$, what implies that the above inequality is always satisfied for any $\alpha_i = 0$.

On the other hand, the second condition holds if, b, the slope of country j's best response function, satisfies

$$0 < \frac{\alpha_j c_i + m}{m} - b\frac{\alpha_i c_j + m}{m}$$
$$\iff b < \frac{\alpha_j c_i + m}{\alpha_i c_j + m}$$

and since the slope of $x_j(x_i)$ is $\alpha_j + m$, we need that

$$\begin{aligned} -(\alpha_j + m) &> -\left(\frac{\alpha_j c_i + m}{\alpha_i c_j + m}\right) \\ (\alpha_j + m) &< \frac{\alpha_j c_i + m}{\alpha_i c_j + m} \end{aligned}$$

and manipulating, and solving for α_i, we obtain the threshold of α_i below which all values of α_i support a zero investment in clean technologies by country i,

$$\alpha_i < \frac{\alpha_j(c_i - m) + m - m^2}{c_j(\alpha_j + m)} = \bar{\alpha}_i(\alpha_j)$$

Case 2: $x_i^* = \frac{\alpha_i c_j + m}{m}$

Let us now analyze the case in which country i sets the maximum investment $\frac{\alpha_i c_j + m}{m}$, while country j does not invest. Firstly, we need that country i's horizontal intercept is above that of country j's, what simply implies

$$\begin{aligned} \frac{\alpha_i c_j + m}{\alpha_i + m} &> \frac{\alpha_j c_i + m}{m} \iff \alpha_i > \frac{\alpha_j c_i m}{m\,(c_j - 1) - \alpha_j c_i} \\ &\iff \alpha_i > \frac{\alpha_j c_i m}{m\,(c_j - 1) - \alpha_j c_i} = \hat{\alpha}_i(\alpha_j) \end{aligned}$$

Secondly, we need that b, the slope of country j's best response function, satisfies

$$0 > \frac{\alpha_j c_i + m}{m} - b\frac{\alpha_i c_j + m}{m}$$

and operating similarly as in the previous case, we have

$$\alpha_i > \frac{\alpha_j(c_i - m) + m - m^2}{c_j(\alpha_j + m)} = \bar{\alpha}_i(\alpha_j)$$

Case 3: $x_i^* = \frac{\alpha_j c_i m + \alpha_i(\alpha_j c_i + m - c_j m)}{\alpha_j m + \alpha_i(\alpha_j + m)}$

Finally, the equilibrium is interior when first, country i's horizontal intercept is below that of country j's, what simply implies,

$$\frac{\alpha_i c_j + m}{\alpha_i + m} > \frac{\alpha_j c_i + m}{m} \Longleftrightarrow \alpha_i < \frac{\alpha_j c_i m}{m\,(c_j - 1) - \alpha_j c_i} = \hat{\alpha}_i(\alpha_j)$$

and second, when b, the slope of country j's best response function, satisfies

$$0 > \frac{\alpha_j c_i + m}{m} + b\frac{\alpha_i c_j + m}{m} \Longleftrightarrow \alpha_i > \frac{\alpha_j(c_i - m) + m - m^2}{c_j(\alpha_j + m)} = \bar{\alpha}_i(\alpha_j)$$

Finally, we must check that $\bar{\alpha}_i(\alpha_j) > \hat{\alpha}_i(\alpha_j)$. Indeed,

$$\bar{\alpha}_i(\alpha_j) - \hat{\alpha}_i(\alpha_j) = \frac{\alpha_j(c_i - m) + m - m^2}{c_j(\alpha_j + m)} - \frac{\alpha_j c_i m}{m\,(c_j - 1) - \alpha_j c_i} > 0$$

for any parameter values. Hence, we can summarize the above three cases as follows:

$$x_i^* = \begin{cases} \frac{\alpha_i c_j + m}{m} \text{ if } \alpha_i > \bar{\alpha}_i(\alpha_j) \\ \frac{\alpha_j c_i m + \alpha_i(\alpha_j c_i + m - c_j m)}{\alpha_j m + \alpha_i(\alpha_j + m)} \text{ if } \alpha_i \in]\hat{\alpha}_i(\alpha_j), \bar{\alpha}_i(\alpha_j)] \\ 0 \text{ if } \alpha_i \in]0, \hat{\alpha}_i(\alpha_j)] \end{cases}$$

where $\hat{\alpha}_i(\alpha_j) = \frac{\alpha_j c_i m}{m(c_j - 1) - \alpha_j c_i}$ and $\bar{\alpha}_i(\alpha_j) = \frac{\alpha_j(c_i - m) + m - m^2}{c_j(\alpha_j + m)}$

4.6.12 *Proof of Lemma 6*

In order to obtain a higher solution than in the case of unconcerned countries (standard public good games), we need to show that the total sum of the optimal investment in clean technologies of the country i and j is greater than one. Notice that there are three possible combinations, which depend of the parameter α_i or α_j.

	x_i	x_j
Case 1	$\alpha_i > \bar{\alpha}_i(\alpha_j)$	$\alpha_j \in]0, \hat{\alpha}_j(\alpha_i)]$
Case 2	$\alpha_i \in]\hat{\alpha}_i(\alpha_j), \bar{\alpha}_i(\alpha_j)]$	$\alpha_j \in]\hat{\alpha}_j(\alpha_i), \bar{\alpha}_j(\alpha_i)]$
Case 3	$\alpha_i \in]0, \hat{\alpha}_i(\alpha_j)]$	$\alpha_j > \bar{\alpha}_j(\alpha_i)$

Case 1 and 3 are trivial because when $\alpha_i > \bar{\alpha}_i(\alpha_j)$ or $\alpha_j > \bar{\alpha}_j(\alpha_i)$ the optimal investment is $x_i^* = 1 + \frac{\alpha_i c_j}{m}$ (or $x_j^* = 1 + \frac{\alpha_j c_i}{m}$) which is greater than one.

In case 2 the optimal investments in clean technologies are $x_i^* = \frac{\alpha_j c_i m + \alpha_i(\alpha_j c_i + m - c_j m)}{\alpha_j m + \alpha_i(\alpha_j + m)}$ and $x_j^* = \frac{\alpha_i c_j m + \alpha_j(\alpha_i c_j + m - c_i m)}{\alpha_i m + \alpha_j(\alpha_i + m)}$, therefore:

$$
\begin{aligned}
x_i^* + x_j^* &> 1 \\
\frac{\alpha_j c_i m + \alpha_i\left(\alpha_j c_i + m - c_j m\right)}{\alpha_j m + \alpha_i\left(\alpha_j + m\right)} + \frac{\alpha_i c_j m + \alpha_j\left(\alpha_i c_j + m - c_i m\right)}{\alpha_i m + \alpha_j\left(\alpha_i + m\right)} &> 1 \\
\frac{\alpha_j m + \alpha_i(m + \alpha_j(c_i + c_j))}{\alpha_j m + \alpha_i(\alpha_j + m)} &> 1 \\
c_i + c_j &> 1
\end{aligned}
$$

Therefore, the total sum of optimal investment in clean technologies is greater than one if and only if $c_i + c_j > 1$.

4.6.13 *Proof of Lemma 7*

Differentiating x_i^* with respect to c_i,

$$
\frac{\partial x_i^*}{\partial c_i} = \begin{cases} 0 \text{ if } \alpha_i > \bar{\alpha}_i(\alpha_j) \\ \frac{\alpha_j(\alpha_i + m)}{\alpha_j m + \alpha_i(\alpha_j + m)} \text{ if } \alpha_i \in]\hat{\alpha}_i(\alpha_j), \bar{\alpha}_i(\alpha_j)] \\ 0 \text{ if } \alpha_i \in]0, \hat{\alpha}_i(\alpha_j)] \end{cases}
$$

which is positive for any parameter values. Similarly, differentiating x_i^* with respect to c_j,

$$\frac{\partial x_i^*}{\partial c_j} = \begin{cases} \frac{\alpha_i}{m} \text{ if } \alpha_i > \bar{\alpha}_i(\alpha_j) \\ -\frac{\alpha_i m}{\alpha_j m + \alpha_i(\alpha_j + m)} \text{ if } \alpha_i \in]\hat{\alpha}_i(\alpha_j), \bar{\alpha}_i(\alpha_j)] \\ 0 \text{ if } \alpha_i \in]0, \hat{\alpha}_i(\alpha_j)] \end{cases}$$

which is negative for $\alpha_i \in]\hat{\alpha}_i(\alpha_j), \bar{\alpha}_i(\alpha_j)]$.

4.6.14 *Proof of Lemma 8*

Differentiating x_i^* with respect to α_i,

$$\frac{\partial x_i^*}{\partial \alpha_i} = \begin{cases} \frac{c_j}{m} \text{ if } \alpha_i > \bar{\alpha}_i(\alpha_j) \\ -\frac{\alpha_j(c_i + c_j - 1)m^2}{[\alpha_j m + \alpha_i(\alpha_j + m)]^2} \text{ if } \alpha_i \in]\hat{\alpha}_i(\alpha_j), \bar{\alpha}_i(\alpha_j)] \\ 0 \text{ if } \alpha_i \in]0, \hat{\alpha}_i(\alpha_j)] \end{cases}$$

which is negative if and only if $c_i + c_j > 1$. Finally, let us differentiate x_i^* with respect to α_j,

$$\frac{\partial x_i^*}{\partial \alpha_j} = \begin{cases} 0 \text{ if } \alpha_i > \bar{\alpha}_i(\alpha_j) \\ \frac{\alpha_i(c_i + c_j - 1)m(\alpha_i + m)}{[\alpha_j m + \alpha_i(\alpha_j + m)]^2} \text{ if } \alpha_i \in]\hat{\alpha}_i(\alpha_j), \bar{\alpha}_i(\alpha_j)] \\ 0 \text{ if } \alpha_i \in]0, \hat{\alpha}_i(\alpha_j)] \end{cases}$$

which is positive if and only if $c_i + c_j > 1$.

4.6.15 *Proof of Proposition 5*

Inserting the results from proposition 1 for two countries with positive weight, $\alpha_i > 0$ and $\alpha_j > 0$, we obtain country i's equilibrium utility level from playing the subgame in which countries invest in emission-reducing technologies.

$$\begin{aligned} U_i &= w - x_i^* + \ln\left[m\left(x_i^* + x_j^*\right) + \alpha_i\left(x_j^* - c_j\right)\right] \\ &= -\frac{\alpha_j c_i m + \alpha_i\left(\alpha_j c_i + m - c_j m\right)}{\alpha_j m + \alpha_i\left(\alpha_j + m\right)} + w - \ln[m] \end{aligned}$$

Since U_i is linear in both c_i and c_j, we can determine the value of c_i and c_j that maximizes U_i by checking if U_i increases or decreases in c_i and c_j. Indeed,

$$\frac{\partial U_i}{\partial c_i} = -\frac{\alpha_i \alpha_j + \alpha_j m}{\alpha_j m + \alpha_i\left(\alpha_j + m\right)}$$

which is negative for all parameter values. In contrast,

$$\frac{\partial U_i}{\partial c_j} = \frac{\alpha_i m}{\alpha_j m + \alpha_i\left(\alpha_j + m\right)}$$

which is positive for all parameter values. Hence, U_i decreases in c_i and increases in c_j.

4.6.16 *Proof of Proposition 6*

In the ultimatum bargaining game country i makes an offer and country j can accept or reject such offer. In the model there are 3 cases, since countries i and j can have different concern levels about green voters. Specifically, country i solves the following maximization problem:

$$\begin{gathered} \underset{c_i, c_j}{Max} U_i(c_i, c_j) \\ s.t. U_j(c_i, c_j) = 0 \end{gathered}$$

First case: Country i's concern level is very high, $\alpha_i > \bar{\alpha}_i(\alpha_j)$, and country j's concern level is small, $\alpha_j \in]0, \hat{\alpha}_j(\alpha_i)]$.

$$\begin{aligned}
(c_i^*, c_j^*) &= \left(\frac{e^{-w}(m - \alpha_i m + e^w \Gamma(\alpha_i^2 + (-1 + \alpha_i + \alpha_i^2)m))}{(-1 + \alpha_i)\alpha_j m}, \frac{\alpha_i(1+m)}{(-1+\alpha_i)}\right) \\
\text{where } \Gamma &= (\alpha_j + m) \\
(x_i^*, x_j^*) &= \left(0, \frac{\alpha_j m c_j + m\Gamma + \sqrt{4\alpha_j^2 \Gamma^2 + m^2(\Gamma - \alpha_j c_j)^2}}{2m\alpha_j \Gamma}\right)
\end{aligned}$$

Second case: Country i and j have a medium concern level, that is $\alpha_i \in]\hat{\alpha}_i(\alpha_j), \bar{\alpha}_i(\alpha_j)]$ and $\alpha_j \in]\hat{\alpha}_j(\alpha_i), \bar{\alpha}_j(\alpha_i)]$.

$$\begin{aligned}
c_i^* &\in \left[0, \frac{\alpha_j w + m(-1+2w) + \lambda \log m}{\alpha_j}\right] \text{ and} \\
c_j^* &\in \left[\frac{-\alpha_j m + \alpha_i m w + \alpha_j \Psi w + (\alpha_j m + \alpha_i \Gamma)\log m}{\alpha_i \Gamma}, \frac{\alpha_i w + m(-1+2w) + \varphi \log m}{\alpha_i}\right] \\
\text{where } \Gamma &= (\alpha_j + m),\ \Psi = (\alpha_i + m),\ \lambda = (\alpha_j + 2m) \text{ and } \varphi = (\alpha_i + 2m) \\
x_i^* &\in \left[\frac{m - mw - m\log m}{\Gamma}, w + \log m\right] \text{ and } x_j^* \in \left[\frac{-m^2 + \Gamma\varphi w + \Gamma\varphi \log m}{\alpha_j m + \alpha_i \Gamma}, w + \log m\right]
\end{aligned}$$

Third case: Country i's concern level is very small, $\alpha_i \in]0, \hat{\alpha}_i(\alpha_j)]$, and country j is very concerned about the noncompliance cost, $\alpha_j > \bar{\alpha}_j(\alpha_i)$.The solution is undefined given that the model's payoff structure is not well-behaved.

Parties Annex I	Non- commitment %	Change 1990-2004	Compliance (1)	Representatives green parties % (2)
Australia	-108	25.1	No	5%
Austria	-92	15.7	No	11%
Belgium	-92	1.4	No	3%
Belarus***	-92	-41.6	45%	2%
Bulgaria* **	-92	-49.0	53%	34%
Canada	-94	26.6	No	0%
Croatia* **	-95	-5.4	6%	0%
Czech Republic* **	-92	-25.0	27%	3%
Denmark	-92	-1.1	1%	3%
Estonia* **	-92	-51.0	55%	1%
European Community	-92	-0.6	1%	5%
Finland	-92	14.5	No	7%
France	-92	-0.8	1%	0.54%
Germany**	-92	-17.2	19%	8.31%
Greece	-92	26.6	No	0%
Hungary* **	-94	-31.8	34%	0%
Iceland**	-110	-5.0	5%	8%
Ireland	-92	23.1	No	0%
Italy	-92	12.1	No	3%
Japan	-94	6.5	No	0%
Latvia* **	-92	-58.8	64%	12%
Liechtenstein	-92	18.5	No	12%
Lithuania* **	-92	-60.4	66%	0
Luxembourg	-92	0.3	No	12%
Monaco	-92	-3.1	3%	0%
Netherlands	-92	2.4	No	5%
New Zealand	-100	21.3	No	3%
Norway	-101	10.3	No	0%
Poland* **	-94	-31.2	33%	0%

Table 3: Green parties and countries' compliance -1st table

Parties Annex I	Non-commitment %	Change 1990 -2004	Compliance (1)	Representatives green parties % (2)
Portugal	-92	41	No	1%
Romania* **	-92	-41	45%	9%
Russian Federation* *	-100	-32	32%	0%
Slovakia* **	-92	-30.4	33%	0%
Slovenia*	-92	-0.8	1%	0%
Spain	-92	49	No	0%
Sweden	-92	-3.5	4%	5%
Switzerland	-92	0.4	No	1%
Turkey ***	-92	72.6	No	0%
Ukraine* **	-100	-55.3	55%	0%
United Kingdom of Great**	-92	-14.3	16%	0%
United States of America	-93	15.8	No	0%

Table 4: Green parties and countries' compliance -2nd table

4.7 APPENDIX OF CHAPTER 2

4.7.1 *Proof of incentive compatibility constraint (2)*

From Branco (1997)

Let us assume that landholder i announces $\widetilde{\theta}_i$ when his true type is θ_i, and all the other landholders truthfully reveal their private types

$$u_i(\widetilde{\theta}_i, \theta_i) = t_i(\widetilde{\theta}_i) - C_i(B_i(\widetilde{\theta}_i), \theta) \tag{a}$$

and therefore the following holds

$$U_i(\theta_i) = u_i(\theta_i, \theta_i) \tag{b}$$

Moreover, we know that a necessary condition for incentive compatibility is

$$\left[\frac{dt_i(\widetilde{\theta}_i)}{d\widetilde{\theta}_i}\right]_{\widetilde{\theta}_i=\theta_i} = \left[\frac{\partial C_i(B_i(\widetilde{\theta}_i), \theta)}{\partial \widetilde{\theta}_i}\right]_{\widetilde{\theta}_i=\theta_i} \tag{c}$$

from (a), (b) and (c), and using the envelope theorem we can have a simple characterization of incentive compatibility as

$$\frac{dU_i(\theta_i)}{d\theta_i} = -\frac{\partial C_i(B_i(\theta), \theta)}{\partial \theta_i} \tag{d}$$

Therefore from the former expression we can write landholder $i\prime s$ utility as

$$U_i^*(\theta_i) = U_i^*(\underline{\theta}) - \int_{\underline{\theta}}^{\theta_i} \frac{C_i(B_i(x), x)}{\partial \theta_i} dx \ \forall \theta_i \in \left[\bar{\theta}, \underline{\theta}\right] \tag{e}$$

4.7.2 *Proof of Proposition 1*

Using (2), the government's problem can be rewritten as

$$\max_{\{B_i(\cdot)\}} \sum_i \int_{\underline{\theta}_1}^{\bar{\theta}_1} .. \int_{\underline{\theta}_n}^{\bar{\theta}_n} \left[\begin{array}{c} V(B_i(\theta_i)) - (1+\lambda)C_i(B_i(\theta_i), \theta) - (1+\lambda)U_i(\underline{\theta}_i) + \\ (1+\lambda)\int_{\underline{\theta}}^{\theta_i} \frac{\partial C_i(B_i(x),x)dx}{\partial \theta_i} \end{array} \right] \times$$

$$\left[\prod_{i=1}^{n} f_i(\theta_i)\right] d\theta_n .....d\theta_1$$

subject to (1) and (3)

If (2) is satisfied then (3) holds if and only if

$$(3\prime)\ U_i(\underline{\theta}_i) \geq (\pi - E_\theta[C_i^M(q, B_{-i}, \theta_i)]),\ \forall i \text{ and } \theta_i$$

Then, the government's objective function can be rewritten as

$$\int_{\underline{\theta}_1}^{\bar{\theta}_1} ..... \int_{\underline{\theta}_n}^{\bar{\theta}_n} [V(B_i(\theta_i)) - (1+\lambda)C_i(B_i(\theta_i), \theta_i) - (1+\lambda)\left[\pi - \int_{\underline{\theta}}^{\theta_i} C_i^M(q, B_{-i}, x_i)dx\right] + \quad \text{(I)}$$

$$(1+\lambda)\int_{\underline{\theta}}^{\theta_i} \frac{\partial C_i(B_i(x),x)dx}{\partial \theta_i}] \times \left[\prod_{i=1}^{n} f_i(\theta_i)\right] d\theta_n .....d\theta_1$$

subject to (1)

and given that $\pi \in \mathbb{R}$ and it is independent of types

$$\sum_i \int_{\underline{\theta}_1}^{\bar{\theta}_1} ..... \int_{\underline{\theta}_n}^{\bar{\theta}_n} \left[\begin{array}{c} V(B_i(\theta_i)) - (1+\lambda)C_i(B_i(\theta_i), \theta_i) + (1+\lambda)\int_{\underline{\theta}}^{\theta_i} C_i^M(q, B_{-i}, x_i)dx \\ +(1+\lambda)\int_{\underline{\theta}}^{\theta_i} \frac{\partial C_i(B_i(x),x)dx}{\partial \theta_i} \end{array} \right] \times$$

$$\left[\prod_{i=1}^{n} f_i(\theta_i)\right] d\theta_n .....d\theta_1 - (1+\lambda)\sum_{i=1}^{n} \pi$$

s.t. (1)

Applying integration by parts to the following elements, we obtain

$$\int_{\underline{\theta}}^{\bar{\theta}} \left[\int_{\underline{\theta}}^{\theta_i} \frac{\partial C_i(B_i(x), x)dx}{\partial \theta_i} \right] \left[\prod_{i=1}^{n} f_i(\theta_i) \right] d\theta_n.....d\theta_1 \quad = \tag{II}$$

$$\int_{\underline{\theta}}^{\bar{\theta}} \frac{C_i(B_i(\theta_i), \theta_i)}{\partial \theta_i} d\theta - \int_{\underline{\theta}}^{\bar{\theta}} \frac{C_i(B_i(\theta_i), \theta_i)}{\partial \theta_i} F(\theta) d\theta \quad = \quad \int_{\underline{\theta}}^{\bar{\theta}} \frac{C_i(B_i(\theta_i), \theta_i)}{\partial \theta_i} \left[1 - F(\theta_i)\right] d\theta_i$$

Thus,

$$\int_{\underline{\theta}}^{\bar{\theta}} \left[\int_{\underline{\theta}}^{\theta_i} C_i^M(q, B_{-i}, x_i) dx \right] \left[\prod_{i=1}^{n} f_i(\theta_i) \right] d\theta_n.....d\theta_1 = \tag{III}$$

$$\int_{\underline{\theta}}^{\bar{\theta}} C_i^M(q, B_{-i}, \theta_i) d\theta - \int_{\underline{\theta}}^{\bar{\theta}} C_i^M(q, B_{-i}, \theta_i) F(\theta) d\theta = \int_{\underline{\theta}}^{\bar{\theta}} C_i^M(q, B_{-i}, \theta_i) \left[1 - F(\theta_i)\right] d\theta_i$$

Substituting (II) and (III) into (I), we obtain the government's relaxed maximization problem

$$\max_{\{B_i(\cdot)\}_{i \neq 0}} \int_{\underline{\theta}}^{\bar{\theta}} \sum_i \left[V(B_i(\theta_i)) - (1+\lambda) C_i(B_i(\theta_i), \theta_i) + \right.$$

$$\left. + (1+\lambda) \times \left(C_i^M(q, B_{-i}, \theta_i) \frac{1 - F(\theta_i)}{f(\theta_i)} + \frac{\partial C_i(B_i(\theta_i), \theta_i)}{\partial \theta_i} \frac{1 - F(\theta_i)}{f(\theta_i)} \right) \right] \times$$
$$\left[\prod_{i=1}^{n} f_i(\theta_i) \right] d\theta_n..d\theta_1 - \sum_{i=1}^{n} \pi$$

subject to (1).

Hence, the first order condition with respect to B_i is

$$\frac{V(B_i(\theta_i))}{dB_i} - (1+\lambda) \frac{C_i(B_i(\theta), \theta)}{\partial B_i} + (1+\lambda) \frac{\partial^2 C_i(B_i(\theta), \theta)}{\partial B_i \partial \theta_i} \frac{1 - F(\theta_i)}{f(\theta_i)} = 0$$

That is,

$$\frac{V(B_i(\theta_i))}{dB_i} = (1+\lambda)\left(\frac{C_i(B_i(\theta),\theta_i)}{\partial B_i} - \frac{1-F(\theta_i)}{f(\theta_i)}\frac{\partial^2 C_i(B_i(\theta_i),\theta_i)}{\partial B_i \partial \theta_i}\right) \tag{IV}$$

Optimal transfer t^*

In order to find the optimal transfer, derive the government's utility, $U_g(B_i) = V(B_i) - t_i^*(B_i)$, with respect to B.

$$\frac{U_g(B_i)}{dB} = \frac{V(B_i)}{dB} - \frac{t_i^*(B_i)}{dB} = 0$$

And since we have that

$$\frac{V(B_i)}{dB} = \frac{t_i^*(B_i)}{dB},$$

Integrating with respect to B and using (IV) we have that:

$$t_i^*(\theta) = (1+\lambda)\left(C_i(B_i(\theta),\theta) - \frac{1-F(\theta_i)}{f(\theta_i)}\frac{\partial C_i(B_i(\theta),\theta)}{\partial \theta}\right)$$

4.8 APPENDIX OF CHAPTER 3

Figures 17-18

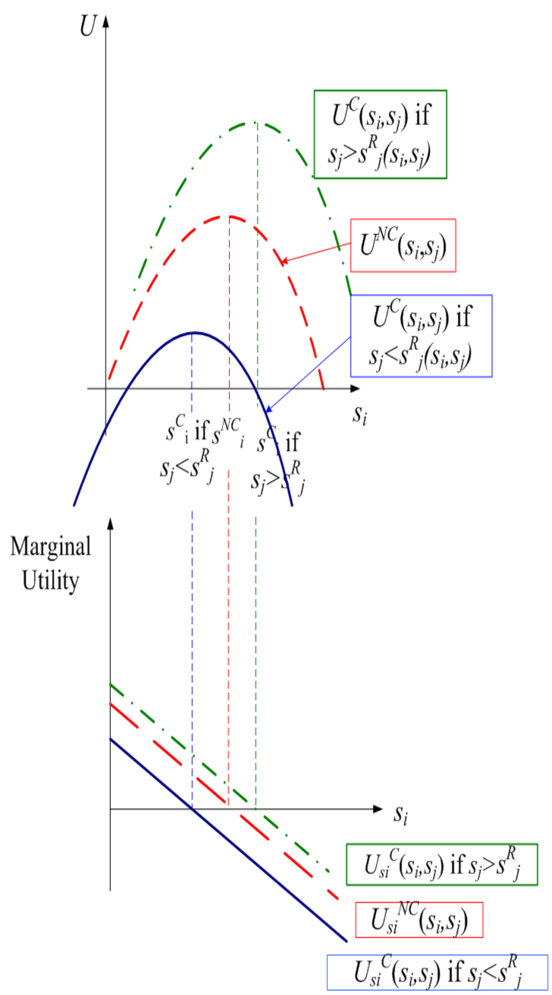

Figure 17: Assumption 5 (a)

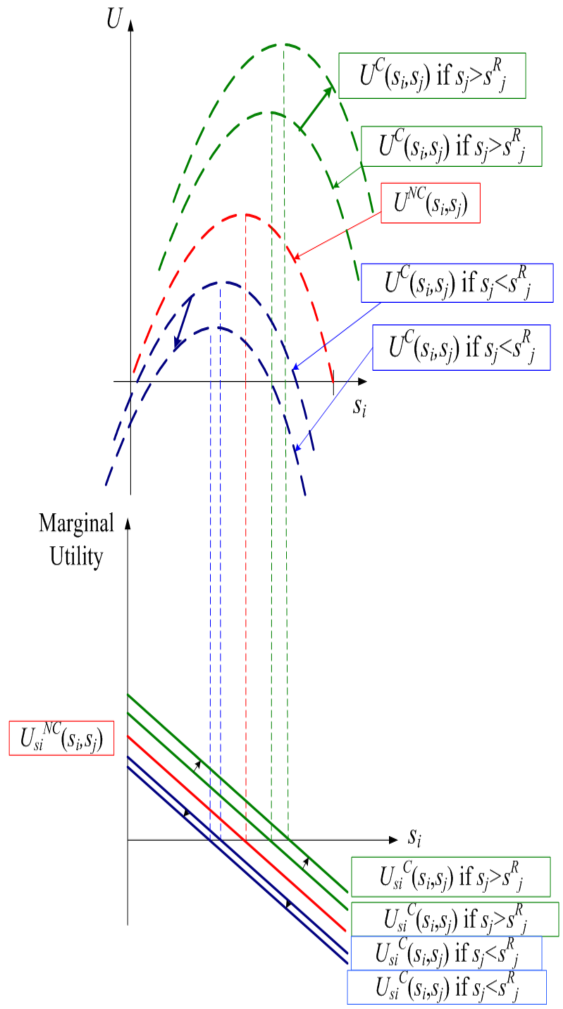

Figure 18: Assumption 5 (b)

4.8.1 *Proof of Lemma 1*

Since $s_i^{NC}(s_j)$ and $s_i^{C}(s_j)$, then

$$\frac{\partial s_i^C\left(s_j\right)}{\partial s_j}=\frac{\partial s_i^{NC}\left(s_j\right)}{\partial s_j}+\left[1-\frac{\partial s_j^{R_i}\left(s_i,s_j\right)}{\partial s_j}\right]$$

where $\frac{\partial D_i(s_i,s_j)}{\partial s_j}=1-\frac{\partial s_j^{R_i}(s_i,s_j)}{\partial s_j}$ given that $D_i\left(s_i,s_j\right)\equiv s_j-s_j^{R_i}\left(s_i,s_j\right)$. Hence, $\frac{\partial s_i^C(s_j)}{\partial s_j}\geq\frac{\partial s_i^{NC}(s_j)}{\partial s_j}$ since $1\geq\frac{\partial s_j^{R_i}(s_i,s_j)}{\partial s_j}$ by definition.

4.8.2 *Proof of Proposition 1*

We first show that player i's best response functions when she is concerned about player j's foregone options and when she is not, respectively, $s_i^C\left(s_j\right)\in\underset{s_i\in S_i}{\arg\max}\,U_i^C\left(s_i,s_j\right)$, and $s_i^{NC}\left(s_j\right)\in\underset{s_i\in S_i}{\arg\max}$ $U_i^{NC}(s_i,s_j)$, contain a single point. Then, we show the result stated in proposition 1.

Note that player i's utility function when she is concerned about player j's unchosen alternatives, $U_i^C\left(s_i,s_j\right)$, is strictly concave in s_i and it is defined over a strictly convex domain. This guarantees that player i's best response function $s_i^C\left(s_j\right)\in\underset{s_i\in S_i}{\arg\max}\,U_i^C\left(s_i,s_j\right)$ contains a single point. A similar argument is also applicable for player i's utility function when she does not assign any relevance to player j's foregone options, $U_i^{NC}(s_i,s_j)$, since it is also strictly concave in s_i and it is defined over a strictly convex domain. Hence, $s_i^{NC}\left(s_j\right)\in\underset{s_i\in S_i}{\arg\max}\,U^{NC}(s_i,s_j)$ also contains a single point.

Once we know that player i's best response function is unique, we just have to compare them in the intervals where $D_i\left(s_i,s_j\right)\geq0$ and $D_i\left(s_i,s_j\right)<0$ in order to check if proposition 1 is satisfied. Let us show it by contradiction. Hence, let us assume that $s_i^C\left(s_j\right)<s_i^{NC}\left(s_j\right)$ when $s_j\geq s_j^{R_i}\left(s_i,s_j\right)$ (i.e. $D_i\left(s_i,s_j\right)\geq0$). Then, for a function $\tilde{s}_i\left(s_j\right)\in S_i$ sufficiently close to $s_i^{NC}\left(s_j\right)$,

$$U_i^C\left(s_i^{NC}\left(s_j\right),s_j\right)-U_i^C\left(\tilde{s}_i\left(s_j\right),s_j\right)\leq U_i^{NC}\left(s_i^{NC}\left(s_j\right),s_j\right)-U_i^{NC}\left(\tilde{s}_i\left(s_j\right),s_j\right)=0$$

That is, player i's marginal utility of raising her strategy when evaluated at the maximizer when she is unconcerned about foregone options, $s_i^{NC}\left(s_j\right)$, is below the marginal utility she could achieve by using this same strategy $s_i^{NC}\left(s_j\right)$ when she is not concerned about player j's unchosen alternatives, which is by definition zero. But this would violate assumption A5 (reciprocity), which states that, when $D_i(s_i, s_j) \geq 0$,

$$U_{s_i}^{C}(s_i, s_j) \geq U_{s_i}^{NC}(s_i, s_j)$$

must hold for any action s_i' sufficiently close to s_i, including $s_i^{NC}(s_j)$. Hence, $s_i^{C}(s_j) < s_i^{NC}(s_j)$ when $D_i(s_i, s_j) \geq 0$ cannot be true. Similarly for $s_i^{C}(s_j) > s_i^{NC}(s_j)$ when $D_i(s_i, s_j) < 0$. Hence, it can only be true that

$$\begin{aligned} s_i^{C}(s_j) &\geq s_i^{NC}(s_j) \text{ for all } D_i(s_i, s_j) > 0 \\ s_i^{C}(s_j) &< s_i^{NC}(s_j) \text{ for all } D_i(s_i, s_j) \leq 0 \end{aligned}$$

4.8.3 *Proof of Lemma 2*

From proposition 1 we know that the difference between player i's best response function when she is concerned and unconcerned about foregone options, $s_i^{C}(s_j) - s_i^{NC}(s_j)$, is weakly increasing in the distance $D_i(s_i, s_j)$. In addition, by assumption A1 we have that player j's utility function $U_j^{NC}(s_j, s_i)$ is strictly increasing in s_i. Hence,

$$U_j^{NC}\left(s_j, s_i^{C}(s_j)\right) - U_j^{NC}\left(s_j, s_i^{NC}(s_j)\right) \text{ is weakly increasing in } D_i(s_i, s_j)$$

Therefore, for two actions $s_j, s_j' \in S_j$ such that $s_j' > s_j$, we have that $D_i\left(s_i, s_j'\right) > D_i(s_i, s_j)$, what implies that

$$U_j^{NC}\left(s_j', s_i^{C}\left(s_j'\right)\right) - U_j^{NC}\left(s_j', s_i^{NC}\left(s_j'\right)\right) \geq U_j^{NC}\left(s_j, s_i^{C}(s_j)\right) - U_j^{NC}\left(s_j, s_i^{NC}(s_j)\right)$$

and rearranging,

$$U_j^{NC}\left(s_j', s_i^{C}(s_j)\right) - U_j^{NC}\left(s_j, s_i^{C}(s_j)\right) \geq U_j^{NC}\left(s_j', s_i^{NC}\left(s_j'\right)\right) - U_j^{NC}\left(s_j, s_i^{NC}(s_j)\right)$$

4.8.4 *Proof of Proposition 2*

Let us s_j^C and s_j^{NC} denote the leader's equilibrium strategies when dealing with a concerned and not concerned follower, respectively. Let us prove $s_j^C > s_j^{NC}$ by contradiction. Hence, assume that $s_j^C < s_j^{NC}$. If this is the case, then the leader's marginal utility from raising her action must be higher when the follower is unconcerned about foregone options than when he assigns a positive importance to them. But this contradicts lemma 2. In particular, recall that lemma 2 states that the marginal utility of raising the proposer's action is higher for the first mover when the second mover is concerned about unchosen alternatives than when he is not. Hence $s_j^C < s_j^{NC}$ must be false, and proposition 2 is satisfied.

4.8.5 *Proof of Proposition 3*

Using Segal and Sobel (1999), we know that player i's preferences over player j's actions can be represented by

$$U_i^C(s_i, s_j) = \gamma_i U_i^{NC}(s_i, s_j) + \gamma_j U_j^{NC}(s_j, s_i) \quad \text{where } \gamma_i, \gamma_j \in \mathbb{R}$$

if preferences satisfy continuity and independence, as well as Segal and Sobel's (1999) condition (★) which states that

$$\text{if } U_i^{NC}(s_i', s_j) = U_i^{NC}(s_i, s_j), \text{ then } s_i' \sim_i s_i \qquad (\bigstar)$$

which are all satisfied in our model.

4.8.6 *Proof of Lemma 3*

Let us consider the receiver's utility function when he assigns a positive importance to the proposer's foregone options and when he does not, respectively, $U_i^C(s_i, s_j, S_j) = x_i + \alpha_i(x_i - f_i)$, and $U_i^{NC}(s_i, s_j) = x_i$. First, assumption A1 is satisfied since $U_i^{NC}(A, s_j') \geq U_i^{NC}(A, s_j')$ and $U_i^{NC}(R, s_j') \geq U_i^{NC}(R, s_j)$ for all s_j since $x_i' > x_i$ if and only if $s_j' > s_j$. Additionally, A2 (concavity) is satisfied since

$$\frac{\partial^2 U_i^C(U_i^{NC}, D_i)}{\partial s_i^2} = \frac{\partial^2 U_i^{NC}(s_i, s_j)}{\partial s_i^2} = 0$$

A3 is trivially satisfied by player i. Regarding assumption A4 (kindness) is satisfied since

$$\begin{aligned}
U_i^C(s_i, s_j) &> U_i^{NC}(s_i, s_j) \text{ since } x_i + \alpha_i(x_i - f_i) > x_i \text{ if } x_i > f_i \\
U_i^C(s_i, s_j) &= U_i^{NC}(s_i, s_j) \text{ since } x_i + \alpha_i(x_i - f_i) = x_i \text{ if } x_i = f_i \\
U_i^C(s_i, s_j) &< U_i^{NC}(s_i, s_j) \text{ since } x_i + \alpha_i(x_i - f_i) < x_i \text{ if } x_i < f_i
\end{aligned}$$

Finally, A5 (reciprocity) is also satisfied, since in their model, $s_i' > s_i$, if and only if $s_i' = A$ and $s_i = R$, what implies that for all $D_i(s_i, s_j) \geq 0$ (i.e., $x_i \geq f_i$)

$$U_i^C(s_i', s_j) - U_i^C(s_i, s_j) \geq U_i^{NC}(s_i', s_j) - U_i^{NC}(s_i, s_j)$$

$$\iff x_i + \alpha_i(x_i - f_i) - 0 \geq x_i - 0 \text{ for any } x_i < f_i$$

and when $D_i(s_i, s_j) < 0$ (i.e., $x_i < f_i$), since

$$[x_i + \alpha_i(x_i - f_i)] - 0 < x_i - 0 \text{ for any } x_i < f_i$$

4.8.7 *Proof of Lemma 4*

Let (x_j, x_i) denote the proposed allocation that the proposer offers to the responder. We know that the responder will accept any offer x_i if and only if the utility he gets by accepting is weakly above than the (zero) utility he gets by rejecting. That is, $x_i + \alpha_i(x_i - f_i) = 0 \iff x_i = \frac{\alpha_i}{1+\alpha_i} f_i$.

Let us now check for sufficiency. Note that the responder does not to accept any offer $x_i < \bar{x}_i$. Instead, accepting any offer $x_i < \bar{x}_i$ would imply negative utility levels, and the responder would be better off by rejecting such an offer, obtaining zero utility. Thus, $x_i < \bar{x}_i$ cannot be an equilibrium strategy.

Finally we need to check that the responder does not reject any offer above $\bar{x}_i$. Let us assume that the responder sets an acceptable threshold $\hat{x}_i > \bar{x}_i$. Then, any offer x_i such that $\bar{x}_i < x_i < \hat{x}_i$ would be rejected, and the responder would find that accepting it constitutes a profitable deviation. Therefore, the acceptance threshold cannot be strictly above $\bar{x}_i$. Hence, the responder does not accept any offer $x_i \in [0, \bar{x}_i)$, but accepts any offer weakly above this threshold level $\bar{x}_i$.

4.8.8 *Proof of Proposition 4*

From lemma 8 we know the responder's acceptance threshold. Since the proposer wants to maximize the remaining portion of the pie which is not offered to the receiver –and guarantees that the receiver accepts such division– he offers $\frac{\alpha_i}{1+\alpha_i}f_i$. This is preferred by the proposer rather than not participating when his remaining share of the pie $1-\frac{\alpha_i}{1+\alpha_i}f_i>0$. That is, the proposer makes the minimal offer $\frac{\alpha_i}{1+\alpha_i}f_i$ if and only if $f_i<\frac{1+\alpha_i}{\alpha_i}$. Since $f_i\in[0,1]$ and $1<\frac{1+\alpha_i}{\alpha_i}$ for any $\alpha_i\geq 0$, then the previous condition $f_i<\frac{1+\alpha_i}{\alpha_i}$ is satisfied for any $\alpha_i\geq 0$. Therefore, the proposer makes the minimal offer $\frac{\alpha_i}{1+\alpha_i}f_i$ for any parameter values.

4.8.9 *Proof of Lemma 5*

Let us consider the worker's utility function when he assigns a positive importance to the proposer's foregone options and when he does not. Respectively, $U_i^C(s_i,s_j)=x_i-e^2+\alpha_i(x_i-f_i)e$ and $U_i^{NC}(s_i,s_j)=x_i-e^2$. Therefore, assumption A1 is satisfied since $U_i^{NC}\left(s_i,s_j'\right)\geq U_i^{NC}(s_i,s_j)$ for any s_i and any $s_j'>s_j$ since $\frac{\partial U_i^{NC}(s_i,s_j)}{\partial s_j}=1\geq 0$. Additionally, A2 (concavity) is also satisfied since

$$\frac{\partial^2 U_i^C(s_i,s_j)}{\partial e^2}=\frac{\partial^2 U_i^C(s_i,s_j)}{\partial e^2}=-2<0$$

A3 is trivially satisfied by player i. Additionally, A4 (kindness) holds since

$$\begin{aligned}
U_i^C(s_i,s_j) &> U_i^{NC}(s_i,s_j) \text{ since } x_i-e^2+\alpha_i(x_i-f_i)e>x_i-e^2 \text{ for any } x_i>f_i\\
U_i^C(s_i,s_j) &= U_i^{NC}(s_i,s_j) \text{ since } x_i-e^2+\alpha_i(x_i-f_i)e=x_i-e^2 \text{ for any } x_i=f_i\\
U_i^C(s_i,s_j) &< U_i^{NC}(s_i,s_j) \text{ since } x_i-e^2+\alpha_i(x_i-f_i)e<x_i-e^2 \text{ for any } x_i<f_i
\end{aligned}$$

On the other hand, A5 (reciprocity) as well since s_i= e and $\frac{\partial U_i^C(s_i,s_j)}{\partial e}=-2e+\alpha\left(x_i-f_i\right)$ and $\frac{\partial U_i^{NC}(s_i,s_j)}{\partial e}=-2e$, then

$$\frac{\partial U_i^C(s_i,s_j)}{\partial e}\geq(<)\frac{\partial U_i^{NC}(s_i,s_j)}{\partial e} \text{ if } D_i(s_i,s_j)\geq(<)\,0$$

4.8.10 *Proof of Lemma 6*

The worker's optimal amount of effort to exert as a function of the wage proposal offered by the firm, $e(x_i)$, can be obtained from solving the following utility maximization problem

$$\max_{e \in \mathbb{R}_+} \; U_i^C(s_i, s_j) = x_i - e^2 + \alpha_i(x_i - f_i(x_i))e$$

Differentiating with respect to e, and manipulating, we have

$$e(x_i) = \begin{cases} \frac{1}{2}\alpha_i (x_i - f_i(x_i)) & \text{if } x_i > f_i(x_i) \\ 0 & \text{otherwise} \end{cases}$$

For sufficiency, just note that the worker will never respond to an offer x_i by exerting a higher effort level than the one specified in $e(x_i)$. Indeed, on the one hand, if he exerts higher effort levels, he will have more disutility from such effort than the utility he derives from the third term of the above utility function for $x_i > f_i(x_i)$. On the other hand, if he exerts less effort, then the marginal utility from exerting additional effort when $x_i > f_i$ (third term of the utility function) would be greater than the marginal disutility from exerting effort (second term). Hence, the worker would be better off by exerting more effort. Hence, the above effort level $e(x_i)$ is optimal for the worker when the wage offered is x_i.

4.8.11 *Proof of Proposition 5*

As shown in the lemma 2, the worker's optimal effort level is $e(x_i) = \max\left\{\frac{1}{2}\alpha_i (x_i - f_i(x_i)), 0\right\}$.

Regarding the employer offer,we know that the employer inserts the above best response function into his utility function, in order to find the optimal wage offer. $\max\limits_{x_i \in [0,1]} (v - x_i)e(x_i)$. Hence,

$$x_i^* = \frac{v(1 - f_i'(x_i^*)) + f_i(x_i)}{2 - f_i'(x_i^*)} \in \arg\max (v - x_i)e(x_i)$$

Note that the employer prefers to offer $x_i^* = \frac{v\left(1 - f_i'(x_i^*)\right) + f_i(x_i)}{2 - f_i'(x_i^*)}$, where $x_i^* > f_i(x_i^*)$ since $v > 1$ and $f_i'(x_i) < 1$, and induce a positive effort level from the worker, rather than offering any wage level $\hat{x}_i < f_i(\hat{x}_i)$ which induces no effort; see $e(x_i)$. Indeed, the employer's equilibrium utility level from offering x_i^* is $V = (v - x_i^*)\frac{1}{2}\alpha_i (x_i^* - f_i(x_i^*))$, which is positive for any parameter values.

Instead, if the employer makes any offer $\hat{x}_i < f_i(\hat{x}_i)$, the worker exerts no effort, and $U_F = 0$. Hence, x_i^* is indeed the equilibrium wage offer.

Finally, in order to check for the worker's voluntary participation, we need to find what is the minimum offer to be accepted by the worker. That is, we must find a wage offer s_j=x_i such that $U^C(s_i, s_j, S_j) = 0$.

$$x_i - e(x_i)^2 + \alpha_i(x_i - f_i(x_i))e(x_i) = 0$$

$$\Longleftrightarrow x_i - \left(\max\left\{\frac{1}{2}\alpha_i(x_i - f_i(x_i)), 0\right\}\right)^2 + \alpha_i(x_i - f_i)\max\left\{\frac{1}{2}\alpha_i(x_i - f_i(x_i)), 0\right\} = 0$$

In the case in which the foregone option $f_i(x_i) > x_i$, then the above expression is reduced to $x_i = 0$. That is, any wage offer is accepted. On the other hand, in the case in which $f_i(x_i) < x_i$, then, we can reduce the above expression to $x_i = \frac{-2+\alpha_i^2 f_i(x_i)+2\sqrt{1-\alpha_i^2 f_i(x_i)}}{\alpha_i^2}$, which is always negative, for any values of α_i and $f_i(x_i)$. Therefore, the minimum offer to be accepted by the worker in both cases ($x_i > f_i(x_i)$ and $x_i < f_i(x_i)$) will be $\bar{x}_i = 0$, since we are assuming that the firm cannot make any negative offers. Note that in the case that $f_i'(x_i) = 0$ then x_i^* becomes $x_i^* = \frac{v+f_i(x_i)}{2}$.

4.8.12 *Proof of Lemma 7*

Lets us consider two cases, first, the case when the responder's utility function assigns a positive importance to the proposer's foregone options and , second, when he does not, respectively, $U_i^C(s_i, s_j) = z_i + \left[m(s_i + s_j)\left[1 + \alpha_i\left(s_j - s_j^R\right)\right]\right]^{0.5}$ and $U_i^{NC}(s_i, s_j) = z_i + [m(s_i + s_j)]^{0.5}$. Therefore, assumption A1 is satisfied since $U_i^{NC}\left(s_i, s_j'\right) \geq U_i^{NC}(s_i, s_j)$ for any s_i and $s_j > s_j$ given that $\frac{\partial U_i^{NC}(s_i,s_j)}{\partial s_j} = \frac{m}{2[m(s_i+s_j)]^{0.5}} > 0$ for any parameter values.

Similarly, A2 (concavity) is also satisfied since

$$\frac{\partial^2 U_i^C(s_i, s_j)}{\partial s_i^2} = -\frac{m^2}{4\left[m\left(s_i + s_j\right)\right]^{3/2}} \leq 0$$

$$\frac{\partial^2 U_i^C(s_i, s_j)}{\partial s_i^2} = -\frac{m^2(1 + \alpha_i(s_j - s_j^R))^2}{4\left[m(s_i + s_j)\left[1 + \alpha_i(x_i - s_j^R)\right]\right]^{3/2}} \leq 0$$

A3 is trivially satisfied by player i. In addition, A4 (kindness) is satisfied given that

$$\begin{aligned} U_i^C(s_i, s_j) &> U_i^{NC}(s_i, s_j) \text{ for any } s_j > s_j^R \\ U_i^C(s_i, s_j) &= U_i^{NC}(s_i, s_j) \text{ for any } s_j = s_j^R \\ U_i^C(s_i, s_j) &< U_i^{NC}(s_i, s_j) \text{ for any } s_j < s_j^R \end{aligned}$$

On the other hand, A5 holds as well since $\frac{\partial U_i^C(s_i,s_j)}{\partial s_i} = -1 + \frac{m(1+\alpha_i(s_j - s_j^R))}{2\left[m(s_i+s_j)\left[1+\alpha_i(s_j-s_j^R)\right]\right]^{0.5}}$ and $\frac{\partial U_i^{NC}(s_i,s_j)}{\partial s_i} = -1 + \frac{m}{2[m(s_i+s_j)]^{0.5}}$, then it is easy to show that

$$\frac{\partial U_i^C(s_i, s_j)}{\partial s_i} \geq (<) \frac{\partial U_i^{NC}(s_i, s_j)}{\partial s_i} \text{ if } s_j \geq (<) s_j^R$$

4.8.13 *Proof of Lemma 8*

The responder's utility maximization problem is just given by

$$\max_{z_i, G} U_i^C(s_i, s_j) = \max_{z_i, G} z_i + \left[mG\left[1 + \alpha_i\left(s_j - s_j^R\right)\right]\right]^{0.5}$$

$$\begin{aligned} \text{subject to } z_i + s_i &= w_i \\ s_i + s_j &= G \\ s_i, z_i &\geqslant 0 \end{aligned}$$

Using $z_i = w_i - s_i$ and $s_i + s_j = G$ in $U_i^C(s_i, s_j)$, we can reduce the above program to

$$\max_{s_i} U_i^C(s_i, s_j) = \max_{s_i} w_i - s_i + \left[m(s_i + s_j)\left[1 + \alpha_i\left(s_j - s_j^R\right)\right]\right]^{0.5}$$

Differentiating with respect to s_i, and manipulating, we find the best response function for the second mover concerned about the first mover's foregone options.

$$s_i^C(s_j) = \begin{cases} \frac{m\left(1-\alpha_i s_j^R\right)}{4} - \left(1 + \frac{\alpha_i m}{4}\right) s_j \text{ if } s_j \in \left[0, \frac{m\left(1-\alpha_i s_j^R\right)}{4-\alpha_i m}\right) \\ 0 \text{ if } s_j \geq \frac{m\left(1-\alpha_i s_j^R\right)}{4-\alpha_i m} \end{cases}$$

4.8.14 *Proof of Lemma 9*

Regarding the first mover (player i), we know that he inserts the above best response function of the follower into his utility function, $U_j^{NC}(s_j, s_i) = w - s_j + [m(s_i + s_j)]^{0.5}$which is maximized at $s_j^* = \frac{16\left(\alpha_i s_j^R - 1\right) + \alpha_i^2 m^2}{16\alpha_i}$. However, this expression is positive only for sufficiently high values of α_i. In particular, $\frac{16\left(\alpha_i s_j^R - 1\right) + \alpha_i^2 m^2}{16\alpha_i} > 0$ if and only if $\alpha_i > \frac{16}{16 s_j^R + m} = \bar{\alpha}_i$. Therefore,

$$s_j^* = \begin{cases} 0 \text{ if } \alpha_i < \bar{\alpha}_i \\ \frac{16\left(\alpha_i s_j^R - 1\right) + \alpha_i^2 m^2}{16\alpha_i} \text{ otherwise} \end{cases}$$

BIBLIOGRAPHY

[1] Andreoni, James, Paul M. Brown and Lise Vesterlund (2002) "What makes an allocation fair? Some experimental evidence", Games and Economic Behavior, 40, 1-24.

[2] Balboa, A. Daughety and J. Reinganum (2004). "Market Structure and the Demand for Free Trade", Journal of Economics and Management Strategy, 13(1), pp. 125-150

[3] Barrett, S. (1992) "Convention on climate change—Economic aspects of negotiations", Paris, OECD.

[4] Barrett, S. (1994a) "Self-enforcing international environmental agreements", Oxford Economic Papers, 46, pp. 878-94.

[5] Barrett, S. (1994b) "The Biodiversity Suprgame", Environmental and Resource Economics, 4, pp. 111-122.

[6] Barrett, S. (1999) "A theory of full international cooperation", Journal of Theoretical Politics, 11, pp. 519-41

[7] Bereby-Meyer, Yoella and Muriel Niederle (2005) "Fairness in bargaining", Journal of Economic Behavior and Organization, vol. 56, pp. 173-186.

[8] Bolton, Gary E. and Axel Ockenfels (2000) "ERC: A theory of equity, reciprocity, and competition" American Economic Review, 90, 166-93.

[9] Branco, F., 1997. The Design of Multidimensional Auctions. Rand Journal of Economics 28 (1), 63-81.

[10] Brandts, Jordi and Carles Sola (2001) "Reference points and negative reciprocity in simple sequential games", Games and Economic Behavior, 36, 138-57.

[11] Brekke, K., S. Kverndokk and K. Nyborg (2003) "An Econoic Model of Moral Motivation", Journal of Public Economics, 87, pp. 1967-1983.

[12] Camerer, Colin (2003) *Behavioral Game Theory*, Princeton University Press.

[13] Cason, T., Gangadharan, L., Duke, C., 2003. A Laboratory Study of Auctions for Reducing Non-Point Source Pollution. Journal of Environmental Economics and Management 46 (3), 446-471.

[14] Cesar, H. (1994) *Control and game models of the greenhouse effect*. Berlin: Springer-Verlag.

[15] Charness, Gary and Matthew Rabin (2001) "Understanding social preferences with simple tests" Department of Economics Working paper No. E02-317, University of California-Berkeley.

[16] Chayes, A., and A. Chayes. (1995) "The New Sovereignty: Compliance with International Regulatory Agreements", Cambridge, MA: Harvard University Press.

[17] Che, J., 1993. Design Competition through Multidimensional Auctions. RAND Journal of Economics 24 (4), 668-680.

[18] Cox, James C. (2001) "On the economics of reciprocity", University of Arizona, working paper.

[19] Cox, James C. Daniel Friedman and Steven Gjerstad (2005) "A tractable model of reciprocity and fairness", University of Arizona, working paper.

[20] Cox, James C., Klarita Sadiraj and Vjollca Sadiraj (2002) "Trust, Fear, Reciprocity, and Altruism", University of Arizona, working paper.

[21] Cox, James C. and Cary A. Deck (2002) "On the nature of reciprocal motives", University of Arizona, working paper.

[22] Cox, James C. (2003) "Trust and Reciprocity: implications of game triads and social contexts", University of Arizona, working paper.

[23] Dasgupta, S., Spulber, D.F., 1990. Managing Procurement Auctions. Information Economics and Policy 4 (1), 5-29.

[24] Downs, G., D. Rocke and P. Barsoom (1996) "Is the Good News about Compliance Good News about Cooperation?", International Organization, 50(3), pp. 379–406.

[25] Espinola-Arredondo, A. and F. Munoz-Garcia (2007) "The importance of foregone options", University of Pittsburgh, working paper.

[26] Falk, Armin., Ernst Fehr and Urs Fischbacher (2003) "On the nature of fair behavior", Economic Inquiry, 41(1), 20-26.

[27] Fehr, Ernst and Simon Gachter (2000) "Fairness and retaliation: the economics of reciprocity", Journal of Economic Perspectives, 14, 159-81.

[28] Fehr, Ernst and Klaus Schmidt (1999) "A theory of fairness, competition and cooperation", Quarterly Journal of Economics, pp. 817-68.

[29] Fehr, Ernst, Georg Kirchsteiger and Arno Reidl (1993) "Does fairness prevent market clearing? An experimental investigation", Quarterly Journal of Economics, 108, 437-60.

[30] Folmer, H., P. van Mouche and S. Ragland (1993) "Interconnected games and international environmental problems", Environmental and Resource Economics, 3, pp. 313-35.

[31] Francis, W., L. Kenny, R. Morton and A. Schmidt (1994). "Retrospective Voting and political Mobility", American Journal of Political Science, 38(4), pp. 999-1024

[32] Fudenberg, D. and J. Tirole. (1984). "The Fat-Cat effect, the Puppy-Dog Ploy, and the Lean and Hungry Look", American Economic Review, 74(2), pp. 361-366

[33] Guth, Werner, Rolf Schmittberger, and Bernd Schwarze (1982) "An experimental analysis on ultimatum bargaining", Journal of Economic Behavior and Organization, 3: 367-88.

[34] Hanley, N. and H. Folmer (1998) *Game Theory and the Environment*. Edward Elgar publishing.

[35] Hansen, R.G., 1988. Auctions with Endogenous Quantity. RAND Journal of Economics 19 (1), 44-58.

[36] Hoel, M. and K. Schneider (1997) "Incentives to participate in an international environmental agreement", Environmental and Resource Economics, 9, pp. 153–170.

[37] Jehiel, P., Moldovanu, B., Stacchetti, E., 1996. How (not) to Sell Nuclear Weapons. American Economic Review 86 (4), 814-829.

[38] Jehiel, P., Moldovanu, B., Stacchetti, E. , 1999. Multidimensional Mechanism Design for Auctions with Externalities. Journal of Economic Theory 85 (2), 258-294.

[39] Jeppesen, T. and P. Andersen (1998). "Commitment and fairness in environmental games". In Hanley, N. and H. Folmer (eds.), *Game Theory and the Environment*, Edward Elgar publishing.

[40] Kaszegi, Botond and Mathew Rabin (2006) "A model of reference-dependent preferences", Quarterly Journal of Economics, forthcoming.

[41] Klemperer, P., 2002. What Really Matters in Auction Design. Journal of Economic Perspectives 16 (1), 169-190.

[42] Laffont, J., Tirole, J., 1986. Using Cost Observations to Regulate Firms. Journal of Political Economy 94 (3), 614-641.

[43] Laffont, J., Tirole, J., 1993. A Theory of Incentives in Procurement and Regulation. MIT Press. Cambridge, Massachusetts.

[44] Latacz-Lohmann, U., Van der Hamsvoort, C., 1997. Auctioning Conservation Contracts: A Theoretical Analysis and Application. American Journal of Agricultural Economics 79 (2), 407-418.

[45] Latacz-Lohmann, U., Van der Hamsvoort, C., 1998. Auctions as a Means of Creating a Market for Public Goods from Agriculture. Journal of Agricultural Economics 49 (3), 334-345.

[46] Loomes, Graham and Robert Sugden (1982) "Regret theory: an alternative theory of rational choice under uncertainty", Economic Journal, 92: 805-24.

[47] Mas, Alexandre (2006) "Pay, reference points and police performance", Quarterly Journal of Economics, August, vol. 121, no. 3, pp. 783-821.

[48] McAfee, P.R., McMillan, J., 1987. Auction and Bidding. Journal of Economic Literature 25 (2), 699-738.

[49] Nyborg K. (2000). "Homo Economicus and Homo Politicus: Interpretaion and Aggregation of Environmental Values", Jounal of Economic Behavior and Organization, 42, pp. 305-322.

[50] Osborne, M. and A. Rubinstein (1994). *A Course in Game Theory*. MIT Press, Cambridge, Massachussets.

[51] Rabin, M. (1993) "Incorporating fairness into game theory and economics", American Economic Review, 83(5), pp. 1281-1302.

[52] Riordan M., Sappington, D., 1987. Awarding Monopoly Franchises. American Economic Review 77 (3), 375-387.

[53] Rubinstein, A. (1982) "Perfect Equilibrium in a Bargaining Model", Econometrica, 50, pp. 97-110

[54] Rubio, S and A. Ulph. (2007) "An Infinite-horizon Model of Dynamic Membership of International Environmental Agreements", Forthcoming in Journal of Environmental Economics and Management.

[55] Schelling, T. (2006). *Strategies of Commitment and Other* Essays. Harvard University Press, Cambridge, Massachussets.

[56] Skreta, V., Figueroa, N., 2004. Optimal Auction Design For Multiple Objects with Externalities. Working paper, University of California at Los Angeles.

[57] Stoneham, G., Chaudhri,V., Strappazzon, L., Ha, A., 2003. Auctions for conservation contracts: an empirical examination of Victoria' BushTender Trial. The Australian Journal of Agricultural and Resource Economics 47 (4), 477-500.

[58] WHALLEY, J. (1991) "The interface between environmental and trade policies", The Economic Journal, 101, pp. 180-9.