Essays on Political and Environmental Economics

Lluís Puig-Gonzalez

Doctor of Philosophy

University of York

Economics and Related Studies

March 2023

Abstract

This thesis presents three essays, where Chapter 1 is based on political economics, and Chapter 2 and 3 combines both political and environmental economics. Chapter 1 examines trust in the national parliament, using a model to understand factors impacting political decision-making. It starts with electoral competition where parties competes for votes. Once in power, the party receives two sets of information: the new state-probability (internal pressure) and lobbying proposals (external pressure). All these elements will shape the potential final implemented policy. Additionally, trust in national parliaments will be then determined based on the final policy's expected outcomes compared to electoral promises.

Chapter 2 assesses the effectiveness of raising awareness about renewable energy as a climate change mitigation strategy. It employs a two-period model within the energy market, and the study explores various market conditions. Results indicate that when fossil fuel reserves are not fully depleted, increased green consumer preferences can enhance environmental well-being, but this effect weakens if it promotes new renewable energy use, potentially leading to the "weak Green Paradox." Additionally, low-cost renewable energy may extend emissions growth, causing a "strong Green Paradox." In scenarios with exhausted fossil fuel reserves, anticipation in extraction levels occurs increased green consumer preferences. The chapter examines the impact of market structures and other factors on environmental deterioration.

Chapter 3 examines COP26, specifically the key debate centered on choosing between "Phase Down" or "Phase Out" strategies for unabated coal power. Using a three-stage agreement formation game, countries initially decide whether to join the coalition, influencing their pro-environmental commitment. In the second stage, signatories cooperate, while non-signatories make independent resource allocation decisions. In the final stage, countries individually determine energy production levels. Pro-environmental behavior motivates coalition participation. "Phase Down" strategies are more vulnerable with lower pro-environmental behavior, while higher levels effectively curb pollutant extraction.

Contents

List of Tables

List of Figures

Acknowledgements

There is nothing more rewarding than the experience obtained during this doctoral thesis. Nevertheless, the achievement of this goal would not have been possible without the support of the group of people that helped me during this journey. First, I would like to thank *Dr. Bipasa Datta* for all her academic and emotional support, encouragement, advice, and comments received in the last years.

I am also thankful to my thesis advisory panel members, *Dr. Andrew Pickering*, *Mr. John Bone*, and *Dr. Giacomo De Luca*, for their constructive comments, insightful feedback, and helpful discussions relevant to my future academic career. I would also like to thank all the administrative staff at the Department of Economics and Related Studies and the University of York for the help provided in my studies.

Special thanks to the Department of Environment and Geography for considering me in all their academic projects, having crucial support from *Dr. Simone Martino*, *Dr. Marco Sakai* and *Dr. Julia Touza* on the achievement of my goals. To my parents, *Alicia and Luis*; siblings, *Laia and Marc*; sister-in-law, *Maite*; nephew,*Nil*; aunt, *Cristina*; and all my extended family, thanks for all the unconditional support received in the last years. They constantly encourage me to try new challenges and achieve my goals.

I am grateful for all the academic staff and colleagues that have shared time with me on this journey, particularly *Akseer*, *Roberto*, *Alejandro* and *Willis* for making this period memorable and enjoyable. Additionally, I would like to address some words to those friends that were with me in this process.I am grateful for their emotional support, especially, those friends that did not ask to return 36 cents; those friends that always made unexpected calls during the day; those friends that have shared some academic or professional moments with me; those that, despite living abroad, make me feel at home and those friends that have made a simple coffee time a great learning experience. They were always supporting my dreams and helping me to face the bad moments.

Finally, I would like to thank *YOU*, the reader of this doctoral thesis. This is a personal valuable piece of work for which I allocate all my knowledge and efforts.

Enjoy it!

Author's Declaration

I declare that this thesis is a presentation of original work, and I am the solo author of the first chapter, and co-author of the second and third with *Dr. Bipasa Datta*. I contributed to every part of the research and *Dr. Bipasa Datta* contributed at various points with revisions and comments. This work has not previously been presented for an award at this, or any other, University. All sources are acknowledged as References.

Lluís Puig-Gonzalez

York, April 2023

Gracias por aquella lucidez, determinación y emoción de tu mirada cuando, a la vez, me dedicabas aquellas palabras que siempre guardaré en mi corazón...

Introduction

This thesis comprises of three chapter, one chapter on political economics and two chapters on political and environmental economics. Chapter 1 contributes to the literature by explaining, through a theoretical approach, the main factors that affect trust in national parliaments. Chapter 2 contributes to the literature on environmental economy by analysing the existence of the Green Paradox in policies that influence the demand-side of energy. Specifically, this chapter is intended to answer whether the influence in the preference of use of renewable energy will reduce the potential environmental damage generated by fossil fuel resources. Chapter 3 provides a new understanding about the outcome obtained in the *Glasgow Climate Pact* in relation to the status of unabated coal power. More precisely, it contributes to analyse under which conditions countries have incentives for either *Phase Down* or *Out* their unabated coal power. Each Chapter is introduced in detail below.

Chapter 1 examines the variation of trust in national parliaments and the main factors that influence their behavior by using a theoretical framework. The role of trust in the political framework is relevant since it influences the electoral turn-out, voter choice, or political efficacy (Cox, 2003, Hooghe and Marien, 2013). There are two different perspectives with respect to the factors that influence political trust. *One* is the performance of the institution and social perspective, as shown by Lewis and Weigert (1985), Rohrschneider (2002) or Hudson (2006). *Other* is the malfeasance decision of political entities as a signal of the poor management of the economy (Obydenkova and Arpino, 2018, Van Erkel and Van der Meer, 2016). Although there is an extended literature that analyses empirically political trust, the theoretical background is still poor.

According to the literature, trust is defined as "*the expected non-negative outcome that an individual receives based on the expected action from another individual or institutions under uncertainty*" in (Bhattacharya et al., 1998). In political terms, trust is based on the lack of commitment of politicians in their electoral propaganda, since it will influence the final expected outcome. The weak commitment of political entities could be explained by using the presence of two factors. First, the difference in state-probability beliefs between the population and politicians. Some papers such as Schultz (1996), Laffont (1999), and Dahm and Porteiro (2008) have contributed to the idea of how parties design their electoral propaganda based on voters' perceptions. Second, the presence of an external agent capable of convincing the politician to implement a different set of policies. The contribution of lobbies in the political outcome is an extended explored area, identifying how they alter bureaucracy (Banerjee, 1997); induce proposals in electoral competition (Grossman and Helpman, 1996); or implement a certain policy (Grossman and Helpman, 1994). However, few characteristics affect the lobby contribution as the culture (Cameron et al., 2009); the ability to pay and refusal power (Svensson, 2003); or the stability of the country (Serra, 2006); among others.

Our model contains two features: the political decision-making process and their potential effect on trust based on expectations of politician state-probability (called as *internal pressure*) and lobby characteristics (called as *external pressure*). We start by designing a sequence of eight stages where the population, parties (and politicians), and lobby interact in order to decide about the final implemented policy. After knowing their own characteristics, the population cast their votes for the most preferred policy designed by the parties. Once a party is selected, it becomes politician and seats in the national parliament. There, it perceives two different shocks. First, a potential change in their state-probability belief, different from the population one. Second, an offer from the lobby where it suggests a potential final implemented policy and the utility transfer. If politician accepts the proposal, then the lobby suggestion becomes the final implemented policy and receives the utility transfer. Otherwise, if the proposal is rejected, then the internal policy, based on the politician state-probability, becomes the final implemented policy. By using some instruments from political economics, to understand the effect of internal pressures, and contract theory, on the influence of external pressures, we conclude that the final implemented policy will be determined as the convex combination between the lobby ideal policy and the politician most preferred policy. Those parameters will be weighted by the own characteristics of the lobby and the political factors (that comprehends population and politician).

Behind previous model construction, individuals compute their trust in national parliaments. As mentioned, individuals are informed about their own, social and politician parameters; however, they do not have precise information about the potential distortion of the state-probability and lobby characteristics. Since trust in national parliaments is based on the formation of the final implemented policy, individuals predict the potential political outcome based on expectations of the state-probability and lobby characteristics. Both pressures, separately and then combined, are analyzed through their effect on individual trust in national parliament.

The results in Chapter 1 show that more differences between states, in the sense of the most preferred policy, will increase the individual trust asymmetry in the population. When external pressures are considered, factors such as the number of individuals, policy discount factor or the politician ratio (that computes the propensity to care about social welfare) increases individual trust. This result comes from the fact that increases the transfer that the lobby needs to do to the politician to convince it to implement the suggested policy. When both pressures are combined, the individual trust will depend on the politician ratio degree. At higher values, individual trust will mostly depend on the *internal pressures*. At lower values, expectations related to the lobby characteristics will mainly affect the individual trust, being all the pressure coming from the *external* one. Our findings suggest that policymakers need to consider such factors when trust in national parliament is considered. The potential existence of a lobby that protects the internal policy designed by the politician will prevent any influence of external agents that persuades their own benefit. Notice that dropping any influence of exogenous factors in the individual trust will affect positively the electoral turn-out, voter choice, or political efficacy.

Chapter 2 examines whether a demand-side policy, such as the promotion of clean energy in consumer preferences, generates the potential consequence of the so-called *Green Paradox* (Sinn, 2008b). The coordination of demand-side policies as a strategy to reduce emissions in favor of less polluted alternatives is a current political movement. For example, the *Avoid-Shift-Improve* is an environmental sustainability approach, used in Germany, to affect the demand for pollutant alternatives by interrupting consumer behavior (Santos et al., 2021). Although the increasing awareness of the use of clean energy is also related to socio-demographic characteristics, (Ramstetter and Habersack, 2020, Tranter, 2011), policy-makers need to be cautious if this initiative does not trigger unexpected negative environmental consequences. As a contribution to the literature, this chapter presents a model that captures the consequences of a demand-side policy that focuses on the preference stimulation of consumers to the consumption of clean resources.

The effectiveness of the climate change policies have been a big question since *Hans-Werner Sinn*, in his academic book Sinn (2012), identified that the design of climate change mitigation programs to prevent emissions could motivate firms to be more partial to extract in current periods rather than future ones. Concerning the emissions generated in the process, Sinn (2008b) argues that the overall level and the timing of greenhouse gas emissions do matter in evaluating the size of climatic change and their cost in the aggregate economy. From here, many papers were providing many examples of climate change policies that were susceptible to the Green Paradox. For example, when the backstop is considered as alternative energy, an increase in the price of fossil fuels could induce the full exhaustion of fossil fuels (van der Ploeg and Withagen, 2012); changes in the extraction cost (Hoel, 2012); announcement of a rising of carbon tax induces fossil fuels owner to extract their resources quickly (Di Maria et al., 2014, Sinn, 2008b); or the likelihood of creating an alternative to fossil fuels (Strand, 2007); among others.

It was not until Gerlagh (2011) that the notion of the Green Paradox was extended. When the effect of low-carbon energy sources and their imperfect substitutes were introduced in the conventional model, Gerlagh (2011) identified two different trends related to the emissions level. Specifically, it was called *weak Green Paradox* as the phenomenon where the climate change policy encourages firms to extract faster fossil fuel resources, raising, at the same time, the current emissions. However, the *strong Green Paradox* occurs when the strategies of policy-makers rush the exploitation of the last unit of fossil fuel resource, increasing the net present value of cumulative damages from global warming. Although there are efforts to comprehend the demand-side, Chakravorty and Krulce (1994), Chakravorty et al. (1997) and Grafton et al. (2012), there is a potential gap to understand better the effect of preferences between fossil fuels and clean energy in the increase in the level of emissions.

Intending to provide some understanding about the effect of addressing consumer preferences to motivate the use of clean energy, we developed a two-periods time model, where pollutant and clean energy are extracted or produced and finally consumed in a market. On the one side, consumers seek to maximize their utility function, which contains a preference parameter that determines the prevalence of clean energy over fossil fuel. As a contribution, we assume a quasilinear function composed of a strictly concave, with a convex combination of both types of energy ruled by the preference parameter, plus a numeraire good. On the other side, there is the multi-supplier firm, newly considered in the literature, that produces both pollutant and renewable energy. Their role is to decide the amount of both types of energy to supply into the market, affecting the resource constraint for future periods. Additionally, we define the competitive equilibrium in our model. This concept will help us to understand the potential effect of the preference parameter in different types of markets as perfect competitive ones or monopoly.

In general, two results are detected. When the characteristics of the economy moves the economy to not fully exhaust their fossil fuel reserves, the exogenous shock generated in consumer preferences contributes positively to environmental welfare, only if it does not induce the incorporation of renewable energy as an alternative source. When the last occurs, the presence of a substitute stimulates the extraction of fossil fuels and rises the current level of emissions, invoking the *weak Green Paradox*. In addition, if clean energy is relatively cheaper, emissions will be extended for the next periods, worsening the net present value of emissions and, consequently, the potential environmental damage, allowing the existence of the *strong Green Paradox*. However, when the characteristics of the economy moves the economy to fully exhaust their fossil fuel reserves, the exogenous shock generated in consumer preferences contributes to an anticipation of the extraction strategies to the current periods, generating more emissions and emerging the *weak Green Paradox*. Since the fossil fuel remains fully exhausted and the environmental damage function is linear without discounting future outcomes, this increase in the preference parameter will not contribute to the *strong Green Paradox*.

When market structures are compared, it generates a diversity of consequences. Following the literature, our model supports the idea that imperfect competitive markets reduce the net present value of emissions compared to the perfect ones, for all the values of the preference parameter. We have observed that perfect competitive markets are the ones that provides more polluting energy in the market, increasing their extraction compared to monopolies and generating, in this case, more emissions. However, if we consider the introduction of clean alternatives, monopoly are the ones who highly reacts, being more susceptible to the strong Green Paradox, if they do not fully exhaust their fossil fuel resources.

Finally, Chapter 2 allows us to have a wide perspective on the consequences of this type of demand-side policy in the environment. Our results suggest some coordination between policies to reduce the potential not tempted impact on the environment. As an illustration, the *United Nations General Assembly* could promote the emergence of the Green Paradox in their *Agenda 2030*. Specifically, the promotion of population education about sustainable development, represented as the 13.2 *Sustainable Development Goals* (or SDG); plus the increase of the affordable renewable share in the total final energy consumption, defined as the 7 SDG; both are potential policies that could contribute in the increase in emissions level. This situation gives us a scenario where an imminent reaction might be considered; because, otherwise, an increasing preference for the use of clean energy (compared to fossil fuels) could motivate the production of new renewable energy methods, stimulating a quick fossil fuel extraction under the presence of an affordable alternative.

Chapter 3 focuses on explaining the main factors and conditions that have conducted the *Glasgow Climate Pact* to determine the *Phase Down* of unabated coal power. In 2021, the city of Glasgow hosted the 26^{th} Conference of Parties (COP26) ruled by the United Nations. After two weeks of debates, the *2021 United Nations Climate Change*

Conference ended up with the *Glasgow Climate Pact*[1], an agreement that identified a plan for the reduction of methane emissions or the decarbonization of road transports, among others. Another important improvement was the recognition of fossil fuels as one of the main drivers of global warming. Since coal represents 46 percent of carbon dioxide emissions worldwide, the COP26 decided to agree on the status of the current unabated coal power among all the 196 countries members of the *2021 United Nations Climate Change Conference*. Although the main intention of the COP26 was to set the *Phase Out* of unabated coal powers, after the influence of some developing countries, such as India and China, the COP26 decided to "*accelerate efforts towards the phase-down of unabated coal power*", with a "*phase-out of inefficient fossil fuel subsidies*"[2]. The COP26 also reunited around USD 350 million from the Adaptation Fund and around USD 600 million from the Least Developed Countries Fund, a weak effort compared to the USD 100 billion a year promised by the developed countries to face the challenges from developing countries.

With these factors in mind, this chapter analyzes the main explanation for the bargaining process captured by the *Glasgow Climate Pact*. Additionally, it attempts to provide an understanding of a new bargaining process not analyzed yet in the literature. The perspective considered in this chapter is to understand the *Glasgow Climate Pact* as an agreement that influences the aggregate emissions level through the resource constraint of country members. This approach is different from the ones observed in the literature since they focus specifically on how the bargaining process affects the structure of the welfare function of countries (the objective function of the optimization problem) rather than the limitation of inputs that their members are able to use. Furthermore, the term pro-environmental behavior (Taufik and Venhoeven, 2018) is internalized in the chapter to represent the potential influence of funding captured by COP26. Such a concept is related to the warm glow defined in the study by Andreoni (1989, 1990), where individuals contribute to a public good because of their altruism and the private benefit obtained from participating. Finally, due to the main intention from the *Glasgow Climate Pact*, the coalition focused on reducing the aggregate damage level, subject to the potential impact that such a decision can generate on the welfare of country members. Mathematically, all the signatory countries will decide the status of the unabated coal power such that minimizes the aggregate damage level, respecting always the participation constraint of those that participate in the agreement.

---

[1] See Glasgow Climate Pact for more information.

[2] From Glasgow Climate Pact, Article 36.

In order to address this research question, this chapter uses a basic linear-quadratic model of an emissions game. Under these characteristics, countries choose the level of dirty (fossil fuel) and clean (renewable) energy to extract or produce, according to their welfare function. It is assumed that countries will differ in their preference parameter related only to dirty energy. Before setting those levels of energy, the country decides whether to join or not an International Environmental Agreement that promotes pro-environmental behavior, even though the status of fossil fuel resources is decided in common. If the country rejects the proposal, it will act as an individual member, where first determines the level of resources to spend and later chooses the level of energy to spend. Otherwise, if the country accepts the proposal, it will enter into a three-stages bargaining process. After the individual decision of joining the coalition, all signatory countries decide unanimously the percentage of fossil fuel resources that they are able to spend in the next stage. Once this decision is in place, the signatory countries will elect individually the level of both types of energy that they will extract or produce, always respecting the decision of the coalition with respect to the maximum amount of input to spend.

As a result, the pro-environmental behavior characterization of the International Environmental Agreement helps the countries to participate in the agreement. Depending on the level of the previous parameter, the coalition will accelerate the efforts towards the *Phase Down* of fossil fuel resources. The *Phase Out* strategy is a potential solution that could bring an increase in the aggregate damage and reduced size of the stable coalition compared to the optimal decision of *Phase Down*. However, large efforts in the pro-environmental behavior will help the International Environmental Agreement to set the *Phase Out* as an optimal solution for the coalition. Apart from the asymmetry of countries with respect to the relevance of fossil fuel in their economies, it is also important to observe that COP26 was ineffective to provide the financial support required to halt the extraction of the fossil fuel reserves to reduce the aggregate damage.

Chapter 1

Walk the talk.

Pressures effects in policy decision-making and their influence on trust in national parliaments.

(Solo-authored chapter)

1.1 Motivation

The concept of trust has become important in the political framework, given its influence on electoral turn-out, voter choice, or political efficacy, as specified in Cox (2003) or Hooghe and Marien (2013), among others. Further research has been developed around the main covariates that influence the dynamic of political trust, motivated by the issues of power, public behavior, or protests observed over recent years. On the one hand, the performance of the institution and social perspective are identified with high explanatory power concerning the trust in the political institution, as analyzed in Lewis and Weigert (1985), Rohrschneider (2002) or Hudson (2006). On the other hand, the malfeasance decision of political entities contributes to the variation of this typology of trust, since it provides a signal of the poor management of the economy (Obydenkova and Arpino, 2018, Van Erkel and Van der Meer, 2016). Nevertheless, given the empirical studies that identify the main determinants for the formation of trust in the political institution, it is lacking a potential theoretical approach that captures those ideas. With this purpose, along the following lines, we will provide a perspective of how trust is constituted and which are the potential factors that could determine it.

Conceptually, trust is often related to honesty, confidence, or belief in a certain outcome. In the political framework, we can still use this concept if we contemplate elections, where parties (and future politicians) announce the policies that they would apply in the national parliament. Assuming that individuals rely on this proposal, they could expect the final implemented policy in the next legislature, predicting their pay-off at the end of this period. But, will politicians respect the electoral proposal? If it is so, individuals will trust fully in the policies revealed in the election. However, any political deviation will affect the potential political outcome, having some implications on the perspective of individuals. As a result, there will be a lack of confidence in the political decision. If it occurs, will individuals react equally under any final implemented policy different from the electoral one?

Under the several reasons behind the non-full commitment in the electoral word, we will distinguish between two different factors: the *internal* and *external pressures*. The former occurs when a politician does not share the same information as the population, since they could update their current beliefs, such as imminent crisis or pandemics, being able to change their policies from the ones established in elections. From this asymmetric information scenario, individuals could expect a different outcome from the one revealed in elections. The latter occurs when the politician decision making-process is influenced by a third agent, called *lobby*. In this typology of pressure, a politician receives a policy proposal, which it must either accept or not. If its accepted, the lobby proposal affects the political outcome, being able to reduce the expected politician utility at the end of the legislature. In this case, the lobby will provide a transfer as a counterpart to help the politician to accept their offer. As an illustration, corruption could be a potential example of *external pressures* when the lobby proposes a policy in exchange for economic aid. Another example could be international institutions orders, where their implementation prevents the government to be punished. Finally, the presence of both pressures will affect the trust in the national parliament, since it could imply a change in the political outcome obtained at the end of the legislature.

According to this perspective, we aim to design a theoretical model that could help us to understand the identification problem of the potential factors that could influence individual political confidence, being able to assess the consequences of previous pressures on individual trust. As required, the political-decision making will be designed in a sequence that allows the interaction of all players: the population, parties (or politician), and lobbies. After identifying the equilibrium, trust will be defined and computed as the individual likelihood of obtaining a higher pay-off than the one promised under an electoral competition. Once defined, both *internal* and *external pressures* will be under review, concluding how they influence the trust in national parliaments.

In order to achieve this goal, we will present our research as follows. First, in section 1.2, we will mention the literature as an approach to building our model. Next, we will proceed by introducing our model in section 1.3, where the timing and the agents characteristics are defined. The previous section will help us to determine the equilibrium in our model in section 1.4. Once we have identified this result, we will compute the trust function in section 1.5, in order to analyze the future results in section 1.6 and conclude with some remarks in section 1.7.

1.2 Literature Review

Our purpose is to understand, through a theoretical economic model, how the concept of trust in national parliaments is formed. Before starting with the relevant literature about our model, it is important to clarify the concept of trust.

Among all the social science fields, there is a joint basis that encompasses trust as an individual feature that captures the belief in exchanges between people, see Lewicki and Bunker (1995) and Cook and Santana (2018). It is believed that the formation of trust depends on many different factors: as past experience (Glanville and Paxton, 2007, Uslaner, 2002); or network effects (Glanville, 2016); among others. Both of them will contribute to the formation of trust explained in Bhattacharya et al. (1998), where it is defined as "*the expected non-negative outcome that an individual receives based on the expected action from another individual or institutions under uncertainty*". In mathematical words, they compute the probability that an individual gets a pay-off weakly higher than the one defined previously with a second player. Due to this interaction, they formalized trust in three different scenarios: simultaneous, sequential, and deterministic outcomes. Thanks to this approach, we interpret trust in the national parliament as the belief of individuals to get, at least, the same outcome that they had arranged with a politician in an electoral competition. In other words, we will understand that in elections, both politician and individuals decide their best strategies, by designing their electoral propaganda or assigning their votes, respectively. At this level, both will determine their expected pay-off level at the end of the legislature. But, given the politician's power of modifying policies, there is a likelihood that the final implemented policy will be different from the electoral one. Hence, the trust could be diminished if any political modification reduces the final outcome received by individuals. However, regardless of where their behavior comes from, it is proved that political trust is linked with the turn-out (Cox, 2003, Grönlund and Setälä, 2007); the voter choice (Bakker et al., 2016, Hetherington and Husser, 2012); and the political efficacy (Hooghe and Marien, 2013). In the last one, they show that positive effects in trust conduct to the reduction of non-institutionalized participation (such as the petition, the boycotted products, or the demonstrations) organized by non-elite actors. As we may induce, both three consequences interfere with the national eligibility criteria and affect the welfare benefits generated by democracy, (Larsen, 2007).

After defining trust, we will focus on the different reasons that conduct a political non-commitment. Before starting, it is crucial to understand how electoral promises are settled. In Palfrey (1984), it identifies the electoral competitive equilibrium in a framework where vote-maximizing parties choose their optimal policy. In addition, it considers that their strategies are rationally anticipating the entrance of a third party. Close to this idea, Feddersen et al. (1990) developed a timing where candidates decide whether to enter or not in the electoral competition. After this stage, those that have entered choose their optimal utility policy over the expected policy outcomes. From this game, they identified that the equilibrium in the electoral competition will be located in the median voter's policy. Those conclusions will help us to determine the electoral competitive equilibrium. However, we need to consider two different factors that would drive the politician to deviate from the electoral outcome: changes in state-probability belief and lobby contributions.

Starting with the effect of state-probability, many references claim their relevance in the politician's decision-making. In Schultz (1996), it considers that electoral candidates are the exclusive agent who knows about the real state of the world. But, under the potential divergence in beliefs, the electoral candidates will consider population intuitive criteria in order to design their policies across states. In Laffont (1999), they analyze asymmetric information between policy maker and individuals, generating the existence of interest groups that participates in the policy implementation by offering bribes, campaign contributions, and other favors. Similarly, Dahm and Porteiro (2008) focus also on the political competition where parties anticipate the most likely state in the economy. In addition, they include a lobby that aims to implement their most preferred policy by investing and offering information about the real state of the economy to the politician. From previous papers, we will assume a potential bias of the politician's state-probability belief concerning the population one. But, in our model, we will exclude any potential lobby influence in the new politician's belief, being this knowledge exogenous from the model. As we will see later, any difference between individuals' and politicians' state-probability will generate a deviation from the implemented policy compared to the promised one.

For the remaining factor, lobby contributions, there are many different ways to influence the political outcome in an economy, such as affecting information about states (Dahm and Porteiro, 2008); altering bureaucracy (Banerjee, 1997); inducing proposals in electoral competition (Grossman and Helpman, 1996); or implementing a certain policy (Grossman and Helpman, 1994). Despite the large range of lobby mechanisms to interrupt the political outcome, we will reduce our perspective by not distinguishing among them. In our case, we will specify lobby contribution as the contract that contains the lobby policy and the respective transfer. Although a politician could receive many lobby offers, we will restrict it to accept one of them. For example, Bernheim and Whinston (1986) model this competition in a first bid auction, where every lobby establishes a transfer for every policy level that they would like to implement. There, they define the truthful equilibrium, as the situation where each lobby offers a transfer for the implementation of a policy that reflects their *willingness-to-pay*. From this concept, they show that in every truthful equilibrium, policy maximizes the sum of lobby and politician outcomes, being this level efficient. But, in their analysis, they do not study the possibility that politician chooses a political alternative not designed by lobbies. This situation was analyzed in Kirchsteiger and Prat (2001) by studying the two lobby competition. They identified that, under equilibrium, each lobby suggests their most preferred policy, without politician preference, and the accepted proposal will generate the greater aggregated outcome. It is relevant to consider the current working paper of Duggan (2018), where it largely contributes to the lobby competition in a spatial model with quadratic utilities. There, it is assumed complete information and externalities across lobbies, where the politician preference is determined by the policy level. This model goes in the same direction as the literature, showing that in equilibrium, each lobby proposes the policy that maximizes the aggregate lobby and politician outcome. Additionally, it shows that the equilibrium is constrained efficient[1] under two lobby competitions. The same concept is used under the presence of multiple competitors, where the lobby with the maximal acyclic competitiveness is the winner[2].

From previous models, we observe that the politician preference parameter is important in the bargaining between lobbies. For example, there is a large literature that shows that lobby contributions depend on political factors, such as the individual and/or group past behavior (Tirole, 1996); the culture (Cameron et al., 2009); the personal and country characteristics (Mocan, 2008); the ability to pay and refusal power (Svensson, 2003); the national characteristics as protestant traditions, histories of British rule, developed economies, imports, federal states or degree of democracy, as analyzed in Treisman (2000); or the stability of the country, (Serra, 2006); among others. In our model, we will characterize this politician's characteristics by a parameter that determines the relevance of politicians about their own profit compared to the social welfare.

---

[1] Similar concept in Dekel et al. (2008). The aggregate utility of the winning lobby and politician at the pairwise optimum exceeds the aggregate utility of the competitor lobby and politician at the pairwise optimum

[2] A generalized constraint efficiency result for the comparison between two lobbies

Our approach relies on building a timing to understand the different effects of internal and external pressure on trust. First, an electoral equilibrium policy will be identified as a benchmark scenario, required to be compared to the final implemented policy observed in the last stage of the game. Then, we will observe the effect of both pressures. Starting with the *internal pressure*, it will capture the asymmetric information case between individuals and politicians, where it will occur if the state-probability belief is not equal among these two players. Later, we will introduce the lobby, and their final outcome will depend on the implemented policy. In this situation, a lobby will have the opportunity of suggesting a policy proposal to the politician in exchange for a monetary transfer, contributing to the concept of *external pressure*. Politicians, after observing policy and transfer level, will decide whether to accept or not the proposal. By considering both pressures, the difference between the electoral and final implemented policy will emerge, being important to determine trust among individuals. In the end, we will run simulations to observe the effect of both types of pressures on trust in national parliaments.

1.3 Model

Consider a population of N individuals characterized by their own political preferences. They are able to allocate a vote in order to choose who is going to seat in the parliament for the next legislature. In addition, we consider the presence of parties, with the aim of becoming politicians in order to determine their own set of policies, and lobbies, whose profit depends on the final policy implemented. All players are affected by the policy level finally implemented in a legislature, $P = \{p \in \mathbb{R} \mid 0 \leq p \leq 1\}$. We will assume that individuals are not sure about the real state of the economy in the next legislature, but they believe that the economy could be in one of the states of the economy $S = \{A, B\}$. Depending on the state of the economy, individuals would prefer a specific policy level.

The decision of players will be restricted to a certain timing, where they interact in order to decide the policy level that will be established in the next legislature. Our timing, specified in Table 1.1 is composed of eight different stages. In the first one, all individuals ($\forall i \in N$) know their preferred policy profile (*political ideology*) and their beliefs about state- probablity. Then, parties, with the previous information, will decide their electoral program (or *promised policy*) and announce them to the population. In the third stage, the population will observe the set of *promised policies* that parties offer, and, depending on their political preference, they will allocate a vote to the closest electoral program to their ideology profile. Therefore, the winner, based on the plurality rule, becomes a politician and seats in the parliament. This role allows the politician to decide the set of policies that will be implemented in the next legislature. In the fifth stage, politician reveals their belief about states. In this case, we will assume that politicians could not share the same beliefs as the population (*internal pressure*). In addition, as the sixth stage, the lobby will have the incentive to persuade the politician to accept their proposal (*external pressure*), the one that would maximize their profit function, in exchange for a payment transfer. Given the offer characteristics, the politician will decide whether to accept or not the lobby suggestion. When a politician decides, the real state is revealed and players receive their own outcome in the last stage.

Stage 1	Nature define individual policy profiles and beliefs about state.
Stage 2	Parties design their promises and announce them.
Stage 3	Population observe the sets of policies and cast votes.
Stage 4	Winner party becomes politician and seats in the national parliament.
Stage 5	Politician reveal their belief about state.
Stage 6	Lobby make their offer.
Stage 7	Politician decide either accepting or rejecting lobby offer.
Stage 8	Real state is revealed and players receive their pay-offs.

Table 1.1: Sequence of the game

Once we have explained the sequence of the game, we will characterize the players in the game. For this reason, we will define players and their roles in the game in the next subsections. In subsection 1.3.1, the population is defined, explaining their utility functions and their role in the model. Later, in subsection 1.3.2, parties are introduced, specifying more their paper in elections and their preferences when they become politician. And, in subsection 1.3.3, it will specify the role of lobbies, with some detail about their profit function and their suggestion method to persuade the politician to modify their final implemented policy.

1.3.1 Population

Population is composed of N individuals, denoted by $i \in (1, ..., n)$, that perceives the impact of a certain policy $p \in P = [0, 1]$ in their utility function. Before starting the legislature, individuals are able to allocate a vote to one of the different parties through an electoral competition.

Consider an individual i who knows its set of preferred policies (or *ideology profile*), $\gamma_i \in [0,1]$, such that, when the policy in the current legislature (p) reaches this level ($p = \gamma_i$), the individual achieves their maximum level of utility ($\bar{u}_i$). When the final implemented policy differs from the preferred policy ($p \neq \gamma_i$), then the individual utility function will be reduced by a quadratic loss multiplied by a parameter $d \in \mathbb{R}^{++}$, named as individual policy discount parameter. Without loss of generality, we will assume that d is common among individuals. According to previous terms, when an individual i faces a potential policy p, their utility level $u_i(p)$ will be:

$$u_i(p) = \bar{u}_i - d(\gamma_i - p)^2 \tag{1.1}$$

where u_i is the utility level function; p is the policy level; $\bar{u}_i$ is the maximum utility level that an individual i may achieve; γ_i is the preferred policy for individual i; and d is the policy discount parameter when policy p do not coincide with the individual i preferred policy γ_i.

From the previous utility function, we observe a quadratic loss function $(\gamma_i - p)^2$ that increases when individual i preference and policy are distant. This quadratic loss function is multiplied by the parameter d that captures the social cost multiplier at any quadratic difference from their most preferred policy. Both discount the maximum utility level $\bar{u}_i$. In this case, we observe that the individual utility function is continuous and *single-peak* at γ_i, increasing the utility function when the policy gets closer to γ_i. In addition, individuals show risk-aversion since $u_i^{''}(p) < 0$.

In our model, we consider that in the next legislature, the entire economy could be in one of both states $S = \{A, B\}$. Depending on it, an individual has different preferences. In this case, we will consider the pair $\{\gamma_i^A, \gamma_i^B\}$ as the level of preferred policies in state A and B, respectively. Furthermore, the probability of state A occurrence will be equal to $\pi_i \in [0,1]$, whereas the remaining probability, $(1 - \pi_i)$, will be allocated to state B. Given this uncertainty, individuals will evaluate policies in their *expected utility function* ($E\left[u_i(p_j)\right]$) that shapes as:

$$E\left[u_i(p)\right] = \pi_i \left[\bar{u}_i - d(\gamma_i^A - p)^2\right] + (1 - \pi_i)\left[\bar{u}_i - d(\gamma_i^B - p)^2\right] \tag{1.2}$$

where $p \in P = [0,1]$ is the policy level; $\pi_i \in [0,1]$ is the individual probability that state

A occurs, whereas $(1 - \pi_i)$ is the individual probability that state B occurs; $\bar{u}_i$ is the maximum utility level that an individual may achieve (unconditional on the state)[3]; the pair of $\{\gamma_i^A, \gamma_i^B\}$ is the preferred policy for individual i in states A and B, respectively; and d captures the policy discount when policy does not coincide with the individual i most preferred policy, unconditional on the state.

Equation 1.2 shows that the individual expected utility function is the convex combination of the individual utility function for each state multiplied by their probability. The individual expected utility function is still single-peak in the individual expected preferred policy across states $(\pi_i \gamma_i^A + (1 - \pi_i)\gamma_i^B)$ and risk-averse since $E\left[u_i(p)\right]'' < 0$. In this case, when p gets closer to the individual expected preferred policy, the individual expected utility function will increase. Concerning the state-probability belief, consider the following assumption:

Assumption 1 *Population is homogeneous in state probability beliefs (π_N):*

$$\pi_1 = \cdots = \pi_n = \pi_N$$

Population will share a common state probability belief π_N. Then, in an electoral competition, parties will design their set of policies considering a common state probability belief π_N. If this statement is not considered, parties will design their set of policies by inducing population state probability belief, competing in order to attract different population groups with similar state probability belief. Then, Assumption 1 restricts parties to compete in policies rather than population beliefs.

When an individual decides where they will cast their vote in an electoral competition, they use Equation 1.2 to compute their preference relationship by satisfying the following property:

Assumption 2 $\forall j, -j \in (1, ..., m)$*, an individual will prefer* $\{j\}$ *over* $\{-j\}$ *if:*

$$E\left[u_i(p_j)\right] \geq E\left[u_i(p_{-j})\right] \Leftrightarrow \{j\} \succsim \{-j\}$$

where $\{j\}$ *and* $\{-j\}$ *denotes party* $\{j\}$ *and* $\{-j\}$*, respectively;* $E\left[u_i(p_j)\right]$ *and* $E\left[u_i(p_{-j})\right]$ *interpret the expected utility level of individual* i *for policies proposed by parties* $\{j\}$ *and* $\{-j\}$*.*

In other words, if a menu of policies from party $\{j\}$ results in a higher individual

---

[3] $\bar{u}_i$ will not indicate the maximum utility level that an individual may achieve, since the convex combination of two independent distribution weighted by two non-zero linear parameters. Theorem 2.3 from Proschan (1965).

expected utility than $\{-j\}$, individual will prefer party $\{j\}$ and $\{-j\}$. Furthermore, if the individual cannot distinguish between policies from party $\{j\}$ and $\{-j\}$ in terms of utility, it will be indifferent between both.

Although all individuals share the same belief with respect to the state-probability, they do differ in their political preferences. As we have revealed before, our economy is composed of a large n individuals that present preferences in policies, that are conditional on the state such that:

Assumption 3 *Population political preferences in any state S is uniformly distributed around $\bar{\gamma}^S = E(\gamma_i^S)$.*

In a nutshell, political preference will be uniformly distributed along political levels (that it will be restricted between 0 and 1, by definition). From this assumption, we observe that the cumulative distribution function of most preferred policies is continuous and strictly increasing on P. This distribution will be centered at $\bar{\gamma}^S$ that it will be denoted as the median most preferred policy in state S. When we add both independent distributions in a convex combination, then:

$$\bar{\gamma}_N = \pi_N(\bar{\gamma}^A) + (1 - \pi_N)(\bar{\gamma}^B)$$

The resulting expected median most preferred policy will be $\bar{\gamma}_N$, where the suffix N indicates the population state-probability belief. By aggregating the utility functions for all n individuals that compose the population, we will find the expected social welfare function, $E[W(p)]$:

$$E[W(p)] = \sum_{i=1}^{n} \left[E\left[u_i(p)\right]\right] = \sum_{i=1}^{n} \left[\pi_N \left[\bar{u}_i - d(\gamma_i^A - p)^2\right] + (1 - \pi_N)\left[\bar{u}_i - d(\gamma_i^B - p)^2\right]\right]$$

Under an egalitarian social economy, where all individuals are considered equal, the resulting expected social welfare function, when policy p is applied, will be equal to $E[W(p)]$. Furthermore, if we consider that $\bar{u}_i$ and d are unconditional on states, our social welfare function will be:

$$E[W(p)] = W_0 - d\sum_{i=1}^{n} \left[\pi_N \left[(\gamma_i^A - p)^2\right] + (1 - \pi_N)\left[(\gamma_i^B - p)^2\right]\right]$$

where $W_0 = \sum_{i=1}^{n} \bar{u}_i$ represents the addition of all maximum utility levels of all individuals[4].

---

[4] This level will not be achieved since individuals are heterogeneous and convex combination of

From previous characteristics, individuals will decide where to allocate their votes in electoral competition. After observing the policies announced by parties, individuals, under a plurality rule with no abstention for one office, can allocate a vote to one party to induce them to seat in the national parliament. In this case, we will assume that individuals allocate *sincerely* a vote to the party whose announced policy reports a higher utility level. Denote v_i as the vector of votes that from an individual i:

$$v_i : (P)^m \rightarrow R = \left\{ (r_1, ..., r_m) : r_j \in \{0,1\} \quad \forall j; \sum_{j=1}^{m} r_j = 1 \right\} \tag{1.3}$$

where v_i will be the vote distribution; and R will be the vector of score ranking r_j and m is the dimension of parties.

Given the policies announced by parties, the population will allocate a unique vote, without abstention, to one party (denoted by $j \in M$, their characterization will be defined later). For the allocation of the vote, individuals will decide *sincerely* where they will cast their vote (known as *sincere vote*):

$$\begin{gathered} \forall \{-j\} \neq \{j\}, \quad u_i(a_j) > u_i(a_{-j}) \Rightarrow r^i_j(p) = r_1 = 1 \\ \forall \{-j\} \neq \{j\}, \quad u_i(a_j) < u_i(a_{-j}) \Rightarrow r^i_j(p) = r_m = 0 \end{gathered}$$

And for all candidates $\{j\}$ and $\{-j\}$:

$$u_i(a_j) \geq u_i(a_{-j}) \Rightarrow r^i_j(p) \geq r^i_{-j}(p)$$

In case where individual is indifferent ($u_i(a_j) = u_i(a_{-j})$), nature will determine $r^i_j(p)$ and $r^i_{-j}(p)$.

A party, with the most preferred policy for individual i, will receive his sincere vote since it is ranked first. In the case where an individual cannot distinguish among them, a fair lottery will decide the ranking. Once everyone assigns their vote, the party with the highest amount of votes becomes politician and seats in the parliament.

Given the sequence of the game, electoral competition does not determine the final implemented policy in a legislature. First, individuals decide where they will allocate their vote among parties' announcement and, once they have decided on the winner, politician decides the policy that they will implement in the legislature. Remember that our model is based on a sequential game where individuals and politician, in an electoral competition, determine a potential strategy that the last player could implement in the future. In this

---

independent states.

case, the politician will decide whether to commit to the announced policy or, depending on the pressures, implement a different one. Individuals are not able to predict the final implemented policy in our game, but they can expect if the announced policy will be applied in the legislature. Under this uncertainty, individuals will determine their trust level as the probability of politician fulfilling their electoral announcement (or *promised policies*).

In order to represent this concept, we will assume that individuals consider the outcome obtained in an electoral competition to set their individual trust, evaluating which policy range will increase their utility level. In other words, individuals determine their trust by assuming which policies will increase their satisfaction in relation to the one obtained in elections. Then, if we consider the promised policies outcome as a benchmark, trust will be defined as:

Definition 1 *Individual trust, $\mathbb{T}_i(\bar{\gamma}_i)$, is a function that computes the probability that the final implemented policy achieves, at least, the promised level of utility obtained in the electoral competition:*

$$\mathbb{T}_i(\bar{\gamma}_i) = Pr\Big(E[u_i(p_F)] \geq E[u_i(p_0)]\Big|\bar{\gamma}_i\Big)$$

where $E[u_i(p_F)]$ is the expected individual utility under the final implemented policy whereas $E[u_i(p_0)]$ is the expected individual utility under the promised policy.

From Definition 1, trust function, $\mathbb{T}_i(\bar{\gamma}_i)$, will provide the likelihood of an individual i, characterized by $\bar{\gamma}_i$, to obtain an expected utility level at the end of the legislature, at least higher than the one obtained from the electoral competition.

1.3.2 Parties and politician

At the beginning of the game, $|M| = 2$ parties compete in an electoral competition for one seat in the national parliament. In this game, parties will design their set of policies depending on their expected political pay-off ($\tau_j(p_j.p_{-j})$):

$$\tau_j(p_j.p_{-j}) = F(p_j|p_{-j}) \tag{1.4}$$

where $\tau_j(p_j.p_{-j})$ is the expected political pay-off that party j will get from announcing $p_j \in P$ over competitor announcement $p_{-j} \in P$. Therefore, $\tau_j(p_j.p_{-j})$ could be interpreted as the individual distribution function that weakly prefers p_j over p_{-j} ($\{j\} \succsim \{-j\}$).

Furthermore, parties have the incentive to participate in the political competition since the minimum utility level achieved in the national parliament exceeds the one obtained by not being elected (normalized at 0), $min_{p \in P} U_P \geq 0$. After announcing their policies, a party will become a politician if their expected political pay-off exceeds the competitor one. In case of a tie, nature decides with the same probability among parties.

When a party becomes a politician and gets a utility function that contains two factors: expected social welfare and own private profit. Both are weighted by a linear parameter, $\beta \in (0,1)$, that measures the importance of politicians about their own profit (whereas $(1-\beta)$ denotes the relevance of politicians about expected social welfare). About this parameter, we will not consider cases where politician utility faces their own private gain ($\beta = 1$), or it is affected purely by the expected social welfare ($\beta = 0$). By considering previous aspects, politician utility function ($U_P(p)$) will shape as follows:

$$U_P(p) = \beta \left[\omega + t(p_L)\right] + (1-\beta)\left[E[W_P(p)]\right] \tag{1.5}$$

where β determines the private gain rate parameter (whereas $(1-\beta)$ captures the magnitude effect of social welfare in politician utility function); ω represents politician wage; $t(p_L)$ is the transfer from the lobby to a politician when the offer is accepted and lobby suggested policy is finally implemented, $(p = p_L)$; and $E[W_P(p)]$ is the social expected welfare level at the set of policies p and politician state probability π_P.

Equation 1.5 show the utility function of a party that seats in the national parliament and becomes a politician. As we have mentioned before, we observe the politician utility function is a convex combination of own private gain (that contains politician wage and lobby transfers) and expected social welfare (that considers all individual's utility under politician state-probability belief). Notice that under no lobby transfers, the politician's utility function will depend on the politician's wage and expected social welfare. But, when the politician agreed in adapting the lobby proposal as a final implemented policy their utility function will include, not only the politician's wage and expected social welfare, but also the transfer made by the lobby.

From the politician utility function, we observe that any policy deviation could generate two shocks. First, the politician could perceive an expected social welfare variation generated by the deviation from promised policies. Any change observed in the expected social welfare is weighted by the parameter $(1-\beta)$. Second, if the deviation of policy is due to a lobby suggestion, the politician could receive an impact in their own private gain function by obtaining the lobby transfer. This impact is weighted by the parameter β in its utility function.

Another aspect to consider from the politician utility is the expected social welfare, $E[W_P(p)]$. Remember that in elections, the politician will design their set of policies under the population state-probability (π_N). But, after becoming a politician, a piece of new information concerning the state-probability is revealed, generating a new belief, π_P. Motivated by this change, the politician will design their set of policies adapting the new state-probability (mentioned as *internal pressures*). Then, their utility function is extended as:

$$U_P(p) = \beta\left[\omega + t(p_L)\right] + (1-\beta)\left[W_0 - d\sum_{i=1}^{n}\left\{\pi_P\left[(\gamma_i^A - p)^2\right] + (1-\pi_P)\left[(\gamma_i^B - p)^2\right]\right\}\right] \tag{1.6}$$

where π_P is the politician probability belief that state A occurs (whereas $(1-\pi_P)$ is the probability belief that state B occurs).

From Equation 1.6, the politician utility function considers the expected social welfare under politician state-probability belief. Then, any policy deviation will generate a quadratic variation discounted by the policy discount parameter d and weighted by the politician state-probability belief π_P.

At stage 5, the politician learns new information about the state-probability (*internal pressures*). As we have seen before, this information changes the politician's perspective with respect to the expected social welfare. Hence, this new information could generate a deviation from the *promised policies* in electoral competition (p_0) to the *internal policy* (p_P) by adapting it in their state-probability belief. After observing this feature, lobbies are able to offer a proposal to the politician in order to adopt a *profitable* policy (*external pressures*) In this case, the politician must decide whether to accept their proposal (and implementing their proposals named as *external policy*, p_L) or reject it and implementing their *internal policy*. In order to persuade the politician, lobbies will design a contract, which contains their political proposal and the transfers that they will do if their offer is accepted. In this case, the lobby must design a beneficial contract to the politician over p_P. In other words, the lobby will attract the politician's acceptance by suggesting the *external policy* that increases their outcome over the one obtained from the *internal policy*. Any contract designed by the lobby will be computed with respect to the *internal policy* level that the politician has determined previously. Then, the politician will accept a lobby proposal if:

$$\underbrace{U_P(p_L)}_{\textit{Accepting contract}} \geq \underbrace{U_P(p_P)}_{\textit{Not accepting contract}} \tag{1.7}$$

where $U_P(p_L)$ indicates the politician utility function when the lobby offer is accepted and implemented; and $U_P(p_P)$ represents the politician utility function when the lobby offer is rejected and the internal policy is implemented.

Equation 1.7 will be satisfied if and only if:

$$t(p_L) \geq \frac{(1-\beta)}{\beta} d \left[\sum_{i=1}^{n} \left\{ (p_L - p_P)^2 + 2(p_L - p_P)(p_P - \bar{\gamma}_{P,i}) \right\} \right] \tag{1.8}$$

where $t(p_L)$ is the transfer made from lobby to politician if lobby suggestion is finally implemented; $\bar{\gamma}_{P,i} = \pi_P \gamma_i^A + (1 - \pi_P)\gamma_i^B$ is the expected individual i preferred set of policies for the pair of $\{\gamma_i^A, \gamma_i^B\}$ in states A and B, respectively; p_P is the internal policy level; and p_L is the lobby suggested policy.

Equation 1.8 shows the transfer range where the politician will accept a lobby proposal. In this case, the politician will implement lobby proposal as a final policy if the transfer weakly exceeds the expected social welfare loss of deviating from internal policies level p_P to the ones offered by the lobby p_L. In other words, the politician will not accept any offer that reduces their pay-off. Hence, it will implement a lobby proposal if the transfer amount covers, at least, the reduction in social welfare generated by implementing p_L. Otherwise, if the previous condition is not satisfied, the politician will not accept a contract from the lobby since their utility function will decrease because of the reduction of the expected social welfare.

1.3.3 Lobby

Our last character, lobby, has the possibility of suggesting a policy $p_L \in P$ to the politician in order to implement it in the next legislature. Given this role, lobbies are aimed to maximize their profit function, which depends, among others, on a certain exogenous parameter ψ, called *lobby maximizer revenue policy level* that optimize their benefit. When the implemented set of policies p is closed to ψ, the lobby increases its profit. Sometimes, the promised set of policies that the politician made in elections p_0 could be far enough from ψ such that lobby would offer a transfer to the politician in order to implement another set of policies close to their maximizer revenue one. In this case, lobbies will have the possibility of offering a new set of policies through a contract that contains the suggested policy (p_L) and the transfer ($t(p_L)$). The last term will be conditional on the implemented policy, in the sense that, if it achieves the suggested policy ($p = p_L$), then the lobby will transfer $t(p_L)$ to the politician. Then, the profit lobby function will be characterized by:

$$\Pi_L(p) = Y - c(\psi - p)^2 - t(p_L) \tag{1.9}$$

where $\Pi_L(p)$ is the lobby profit function when implemented policy is p; Y is the maximum level of revenue that lobby may achieve; ψ is the policy level where lobby maximizes profit; $c \in \mathbb{R}^+$ is the variation in revenue when p is moved towards ψ; p_L is the suggested set of policies from lobby to politician; and $t(p_L)$ is the amount of profit than lobby transfer to politician once the implemented equals the suggested policy, $p = p_L$.

Equation 1.9 shows the lobby profit function depends on two different elements: a quadratic policy loss function and lobby transfers. First, lobbies reduce their profit when implemented policy is away from their maximizer level ψ (*quadratic policy loss function*). In other words, if lobbies can convince a politician to accept their proposal (p_L), then they will increase their profit by moving closer to the implemented policy to their maximizer level. But, doing this action, they will need to compensate the politician for the social welfare losses. In this case, we are introducing the second element that interacts with the lobby profit, lobby transfers. The politician is willing to accept a lobby proposal if they receive compensation for them. As politician will increase their utility function by accepting transfers, lobbies will interact with them through a side-payment in order to implement their suggested policy. Notice that we use the term "*suggested*" set of policies (p_L) instead of maximizer revenue level (ψ), since the lobby must compensate the quadratic social welfare loss to the politician. For this reason, it is not optimal to suggest the maximizer revenue level since it could improve its profit level by proposing a different policy level (close to the promised one). This result is due to the quadratic policy loss function in the expected welfare function. Given this fact, the lobby transfers function $t(p_L)$ is translated as:

$$t(p_L) = \begin{cases} t & \text{if } p = p_L \\ 0 & \text{otherwise} \end{cases} \tag{1.10}$$

where t is the transfer from lobby to politician; p is the policy level; and p_L is the lobby suggested set of policies.

According to Equation 1.10, the lobby transfer will be made if and only if the politician finally implements the suggested policy, p_L. Then, the politician will either accept or reject the offer, by implementing policy p_L. Suppose that the politician decides to accept the offer, but deviate marginally from the policy from the lobby suggested one. In this case, they will not be compensated for moving the policy and, furthermore, they will reduce the expected social welfare, being worse-off than ex-ante, where the set of policies p was modified by internal pressure.

Once we have introduced the lobby behavior, we will understand when they suggest a policy to the politician. Remember that the lobby aims to maximize its own profit. Therefore, the lobby will suggest a policy if it reports a higher profit than the promised ones. In other words:

$$\underbrace{\Pi_L(p_L)}_{\textit{Offering contract}} \quad \geq \quad \underbrace{\Pi_L(p_P)}_{\textit{Not offering contract}} \tag{1.11}$$

where $\Pi_L(p_L)$ indicates the lobby profit when its suggested set of policies is implemented; and $\Pi_L(p_P)$ represents the lobby profit when the politician applies their own set of policies.

Equation 1.11 will be satisfied if and only if:

$$t(p_L) \leq c\left[(p_P - p_L)^2 + 2\,(p_P - p_L)\,(p_L - \psi)\right] \tag{1.12}$$

Equation 1.12 shows the range of potential transfers from the lobby. In this case, the lobby is willing to transfer part of their profit if the transfer level is weakly lower than the revenue variation generated by changing from p_P to p_L. Otherwise, if the previous condition is not satisfied, the lobby would prefer not to suggest a policy to the politician, since the acceptance cost (transfers) is greater than the revenue generated by changing from p_P to p_L.

Then, by considering the politician and lobby transfer range, a lobby proposal will be accepted and implemented if:

$$c\left[(p_P - p_L)^2 + 2\,(p_P - p_L)\,(p_L - \psi)\right] \geq \frac{(1-\beta)}{\beta} d \left[\sum_{i=1}^{n} \left\{(p_L - p_P)^2 + 2(p_L - p_P)(p_P - \bar{\gamma}_{P,i})\right\}\right] \tag{1.13}$$

Equation 1.13 shows the condition for the existence of *external pressure*. A lobby will suggest a policy p_L if the revenue obtained from deviating the policy weakly exceeds the cost of getting the politician's acceptance. Notice that any proposal that satisfies Equation 1.13 will be a core solution since both of them increase their outcome:

$$U_P(p_L, t(p_L)) + \Pi_L(p_L, t(p_L)) \geq U_P(p_P) + \Pi_L(p_P)$$

Once we have described timing and players, we will analyze the procedure of the model in the next section, describing and computing the equilibrium in the different stages of the game. First, we will identify the equilibrium in electoral competition and define the resulting set of promises policies defined as p_0. Later, we will define the equilibrium when the politicians observe another state-probability (p_P) and/or lobby suggests a potential policy (p_L). Both pressures will determine the final implemented policy (p^*).

1.4 Equilibrium

The sequence of the game shows eight stages that explain the potential incentives of politicians to deviate their policies from the original set of promised policies. First, we start by electing the politician among parties in an electoral competition. In order to win, they must offer an interesting set of policies to attract the maximum number of votes. Once they are elected, they receive information about the probability state or lobby suggestions that are characterized exogenously in the game. Remember that our goal is to understand the potential shocks that a politician can receive by seating in the national parliament. Therefore, our model will be composed of two different equilibriums: the electoral and the policy design. In the first one, the parties will determine their set of promised policies. Then, once they are elected, they will receive different incentives to determine their set of implemented policies.

1.4.1 Equilibrium in electoral competition (*Benchmark*)

In an electoral framework where two parties announce their set of policies and individuals cast their votes sincerely over their expected individual utility function, a set of policies and individuals' votes, $\left[\{\bar{p}_j, \bar{p}_{-j}\}, \{\bar{V}_i\}\right] \in P^m \times R^n$, is called electoral equilibrium if:

i) For all parties $j \in M$ and policies $p_j \in P$:

$$\tau_j(\bar{p}_j, \bar{p}_{-j}) \geq \tau_j(p_j, \bar{p}_{-j})$$

ii) For all individuals $i \in N$ and votes $v_i \in R$:

$$E\left[u_i | \{\bar{p}_j, \bar{p}_{-j}\}, (\tilde{v}_i, \tilde{v}_{-i})\right] \geq E\left[u_i | \{\bar{p}_j, \bar{p}_{-j}\}, (v_i, \tilde{v}_{-i})\right]$$

iii) No party or individual uses a weakly dominated strategy.

First, parties will compete through the announcement of their policies to attract more individuals to vote for their platform. Furthermore, individuals will cast their votes for the most preferred political platform. In this case, individuals will decide by evaluating platforms in their individual expected utility function. Given these two factors, parties and individuals will play their non-dominated strategies, converging to the electoral equilibrium.

Proposition 1 *A pair* $\left[\{\bar{p}_j, \bar{p}_{-j}\}, \{\bar{V}_i\}\right] \in P^m \times R^n$ *is the unique two parties electoral equilibrium if and only if* p_0 *coincides with the median voter policy.*
Proof: *See in Appendix, Proof 1.*

Given that individuals are allocated uniformly in the policy space, parties will announce p_0 as their best reply in the electoral competition, since it is a strict Condorcet winner. Moreover, this policy level will coincide with the median voter policy, allocated in $\bar{\gamma}_N$ (named as the social most preferred policy).

One could understand that the equilibrium could be different since individuals have expectations concerning the potential pressures of politicians in their decision-making process. Given individuals decide at the beginning of the game, their expectations will be distributed symmetrically at no existence of any pressure. In other words, the most believed scenario will be that politicians will not receive any pressure to change the final implemented policy from the promised one. But, in the case that a politician receives any pressure (either *internal* or *external*), their potential effects (represented in movements of promised policy level along policy space) will be distributed with the same probability. Given that individuals are risk-neutral, their expected utility function will coincide with Equation 1.2.

From the characteristics of our model, we have computed and identified the equilibrium in the electoral competition (and promised policy) p_0. Therefore, the potential set of policies obtained from this process converges to the median most preferred policy that maximizes the expected social welfare at population state-probability belief. Then, if the finally implemented policy differs from p_0, it will reduce the expected social welfare according to their state probability belief (π_N). In the next subsections, we will determine the equilibrium when a politician decides about the implemented policy under no lobbies, one lobby presence and/or lobby competition.

1.4.2 Equilibrium in national parliament

After defining the equilibrium in electoral competition, the politician seats in the national parliament and receives two sets of information. The first one contains the politician probability-state, which could distort the implemented policy from the promised one ($p_P \neq p_0$). Then, as the second set of information, the politician receives lobby suggestions, where they propose a potential implemented policy p_L. If the politician accepts the proposal, then it receives a transfer, $t(p_L)$. Before proceeding, notice that implementing p_P as a final policy implies a transfer $t(p_P)$ equals 0, since no lobby will pay for remaining at this policy level.

In this subsection, we will analyze different horizons depending on the presence of lobby suggestions. Remember that lobbies will offer a transfer contract to the politician if the aggregate outcome by implementing a p_L (and transferring $t(p_L)$) exceeds weakly the outcome by implementing p_P:

$$U_P(p_L, t(p_L)) + \Pi_L(p_L, t(p_L)) \geq U_P(p_P) + \Pi_L(p_P) \tag{1.14}$$

Equation 1.14 indicates the core solution where politician and lobby get an aggregate outcome weakly higher than the ex-ante one (where *internal policies* were implemented). This situation will generate incentives for lobbies to propose a policy to the politician through a contract. Therefore, two situations will be analyzed. The first one will identify the equilibrium where no lobby suggests a policy level; whereas the second one will be focused on the presence of a lobby in the political decision-making process.

1.4.2.1 Equilibrium under no lobby incentive

Suppose a scenario where the politician decides which policy will implement in the next legislature, since either politician rejects all lobby suggestions or there is not lobby that suggests a policy p_L that satisfies Equation 1.14. In both cases, the politician will decide and implement their set of policies p_P, being in equilibrium if:

$$p_P \in \underset{p \in P}{\operatorname{argmax}} \; U_P(p) = \beta\left[\omega\right] + (1-\beta)E\left[W_P(p)\right]$$

Under the absence of lobby influence, the equilibrium in the implemented policy will be the policy that maximizes the politician's utility. Due to the presence of social welfare in the politician utility function, since $\beta \in (0,1)$, any policy $p \in P \backslash \{p_P\}$ will reduce the social welfare, affecting negatively the politician utility. Therefore, when there is no change between population and politician state-probability, the promised policy will coincide with the finally implemented policy, $p_0 = p_P$. Otherwise, if there is a difference between π_N and π_P, by taking the First Order Conditions of the objective function, we will find that the equilibrium will be:

$$p_P = \pi_P \bar{\gamma}^A + (1-\pi_P)\bar{\gamma}^B \tag{1.15}$$

According to Equation 1.15, under the absence of lobby influence, politicians will choose the expected median most preferred policy weighted by their own criteria about state-probability, π_P. In the next subsections, this equilibrium will be used by lobbies in order to compute their suggestions and transfer that they will offer to the politician. In this sense, the existence of lobbies in our model will depend on the outcome of implementing the *internal policies*.

1.4.2.2 Equilibrium under lobby incentive

Now, suppose that there is a lobby that has incentives to offer a transfer contract to the politician, such that Equation 1.14 is satisfied.

As shown in the timing, the lobby has the chance of offering a transfer contract to the politician. There, the lobby will specify the set of policies that they would implement and the transfer to the politician if their suggestion is finally implemented. After revealing the offer, in the next stage, the politician will decide whether to accept or not. In this case, the politician's decision will depend on the offered outcome. Notice that the politician's profit affects their utility function at parameter β. Then, if the politician will not receive a final outcome weakly higher than the one that they would obtain by implementing p_P, they will not accept the lobby suggestion.

Under the presence of the lobby, this game is split into two different stages, the one when the lobby offers a potential policy p_L and, later, the politician decides whether to accept their proposal or not. Therefore, denote $\sigma^* = (p_L^*, t^*(p_L^*))$ as the sub-game perfect equilibrium if:

i) For lobby:

$$p_L^* \in \underset{p \in P}{\text{argmax}} \;\; \Pi_L(p) = Y - c(\psi - p)^2 - t(p_L)$$

And:

$$\Pi_L(p_L^*, t^*(p_L^*)) \geq \Pi_L(p_P) \Leftrightarrow Y - c(\psi - p_L^*)^2 - t^*(p_L^*) \geq Y - c(\psi - p_P)^2$$

ii) For politician:

$$U_P(p_L^*) \geq U_P(p_P) \Leftrightarrow \beta\left[\omega + t(p_L)\right] + (1-\beta)E[W_P(p_L)] \geq \beta\left[\omega\right] + (1-\beta)E[W_P(p_P)]$$

Therefore, σ^* will be in a sub-game perfect equilibrium if *(i)* the lobby designs an optimal policy that improves its profit with respect to the ex-ante situation (where no suggestion was made); and, *(ii)*, the politician will prefer weakly lobby suggestion (p_L^*) than their own proposal (p_P). Due to the sequence of the game, the lobby knows that the politician will accept any offer that generates an increasing outcome. Then, given the politician participation constraint, the lobby proposes a policy that maximizes their own profit and improves their outcome with respect to the ex-ante situation.

According to the previous definition, a set of strategies $\sigma = (p_L, t(p_L))$ will be in sub-game perfect equilibrium[5] if:

i) The lobby proposes to the politician:

$$\{p_L, t(p_L)\} = \left\{ \frac{c\psi + nd\frac{(1-\beta)}{\beta}\bar{\gamma}_P}{c + nd\frac{(1-\beta)}{\beta}}, \frac{(1-\beta)}{\beta}\left[E_P[W(p_L)] - E_P[W(p_P)]\right]^2 \right\}$$

ii) The politician accepts the lobby proposal.

Starting from the transfers, it will cover losses of implementing the suggested policy p_L instead of p_P. Remember that p_P maximizes the social welfare since it coincides with the median preferred policy at politician state-probability. Then, by construction, any policy p different from p_P will generate a reduction in the expected social welfare since the cost of moving away from the median preferred policy increases. Then, for a certain p_L:

$$E\left[W_P(p_L)\right] \leq E\left[W_P(p_P)\right] \Rightarrow \sum_{i=1}^{n} \left(\bar{\gamma}_{P,i} - p_L\right)^2 \geq \sum_{i=1}^{n} \left(\bar{\gamma}_{P,i} - p_P\right)^2 \Rightarrow t(p_L) \geq 0$$

Any transfer offered to the politician will be non-negative. In this case, the politician will accept a transfer if it is, at least, indifferent between either accepting or rejecting the offer. Under any lobby proposal p_L, politician could reduce their utility, since it is affected by a reduction of expected social welfare at π_P. Then, the lobby needs to compensate for the loss in order to achieve acceptance from the politician, making non-negative the transfer.

With respect to the policy equilibrium, it will be the weighted average between the lobby ideal policy ψ and the social median most preferred policy at the politician probability-state belief $\bar{\gamma}_P$. They will be weighted by the variation in the lobby revenue parameter, c (in case of ψ); and by a multiplicative combination between the politician ratio $\frac{1-\beta}{\beta}$ and the aggregate social policy preference parameters nd, with respect to $\bar{\gamma}_P$. Therefore, when the lobby revenue parameter (c) increase, and it exceeds sufficiently the

---

[5] See Proof 2, for the computation of the sub-game equilibrium

social weight $nd\frac{(1-\beta)}{\beta}$, p_L will tend to the lobby ideal policy:

$$\lim_{c\to\infty} \frac{c\psi + nd\frac{(1-\beta)}{\beta}\bar{\gamma}_P}{c + nd\frac{(1-\beta)}{\beta}} = \psi$$

But, when the social weight increases and exceeds sufficiently the lobby revenue parameter (c), p_L will tend to the median most preferred preference:

$$\lim_{nd\frac{(1-\beta)}{\beta}\to\infty} \frac{c\psi + nd\frac{(1-\beta)}{\beta}\bar{\gamma}_P}{c + nd\frac{(1-\beta)}{\beta}} = \bar{\gamma}_P$$

Then, depending on the weights of lobbies and politician, c and $nd\frac{(1-\beta)}{\beta}$, respectively), the policy in equilibrium will be located in $p_L \in [\psi, \bar{\gamma}_P]$.

Notice that previous result is obtained by binding the politician participation constraint. This result is coming from the lobby stage-dominance, since, starting in an earlier stage, they restrict the transfer at the indifference politician utility level. Obtaining the politician acceptance, the lobby will not increase above this level, since it will reduce their profit. In addition:

Condition 1 *Lobby will offer a contract to the politician if:*

$$nd\frac{(1-\beta)}{\beta} + c \geq 0$$

Proof: *See in Appendix, Proof 3.*

But, by construction, both parameters are weakly greater than 0. When a lobby enters the game, they have incentives to establish p_L. By doing so, the lobby will deviate the policy to their most preferred, allowing them to increase their own profits.

Therefore, we have analyzed the situation where a lobby proposes a policy different from the *internal* one. From the sequence of the game, the lobby proposal will be the weighted average between their maximizer revenue policy level, ψ, and the median most preferred policy at politician state-probability, $\bar{\gamma}_P$. But, in order to get the politician's acceptance, they must offer a transfer equivalent to the losses generated by adapting the lobby proposal as the finally implemented policy. Depending on the revenue characteristics and their most preferred policy level, the difference between *external* and *internal* policy could increase, generating higher transfers from the lobby to the politician. Notice that the unique condition for the existence of this policy is the presence of a lobby, since

Condition 1. From now on, we will consider the final implemented policy p_F as:

$$p_F = \frac{c\psi + nd\frac{(1-\beta)}{\beta}\bar{\gamma}_P}{c + nd\frac{(1-\beta)}{\beta}} \tag{1.16}$$

1.5 Trust in national parliament

In the previous section, the equilibrium in each stage of the game is defined. First, we start by explaining the policy equilibrium in electoral competition with two parties and we have identified that it will be by announcing the median most preferred policy $p_0 = \bar{\gamma}_N$. After being elected, a politician receives the impact of two different types of information: new updates about state-probability, that could differ from the population one ($\pi_N \neq \pi_P$) and lobby suggestions. After observing this information, the politician decides about the new set of policies that it will implement once seats in the national parliament, being able to accept or reject lobby's proposals.

First, we have identified the equilibrium when the politician implements the *internal policy*, described as p_P. As we have shown in Equation 1.15, this policy level is computed by weighting the median preferred policy of each state at the new state-probability ($p_P = \bar{\gamma}_P$). Assume that the difference between the state-probabilities is captured by ϵ. Then, when both the promised and the *internal* policy are compared:

$$p_0 - p_P = \pi_N\bar{\gamma}^A + (1-\pi_N)\bar{\gamma}^B - \left[(\pi_N+\epsilon)\bar{\gamma}^A + (1-\pi_N-\epsilon)\bar{\gamma}^B\right] = \epsilon(\Delta_{\gamma}^{B,A}) \tag{1.17}$$

where $\epsilon \in [-\pi_N, 1-\pi_N]$ is the politician probability-state belief over the individual probability-state belief; and $\Delta_{\gamma}^{B,A} = \bar{\gamma}^B - \bar{\gamma}^A$ indicates the difference between most preferred policies across states.

Notice that both use the median preferred policy of each state weighted at different state-probability. We reduce the difference between both as the difference between state-probability times the distance between social and politician most preferred policy at each state. Given there is no difference among individuals with respect to this last term, internal pressures could distort both equilibriums by assigning a non-zero parameter ϵ.

Later, we have developed the equilibrium when the politician implements the *external policy*. If we do not consider internal pressures ($\pi_N = \pi_P$), the difference between the promised and the external polices will be:

$$p_0 - p_F = \bar{\gamma}_N - \frac{c\psi + nd\frac{(1-\beta)}{\beta}\bar{\gamma}_N}{c + nd\frac{(1-\beta)}{\beta}} = \frac{c(\bar{\gamma}_N - \psi)}{c + nd\frac{(1-\beta)}{\beta}} \tag{1.18}$$

Notice that the lobby characteristics influence the difference between both policies. Specifically, lobby maximizer profit policy (ψ) and variation revenue (c) sizes in the distortion of external pressures in the promised policy. For example, the existence of distortion is generated by the maximization problem of the lobby at a policy not equal to the median most preferred one, being magnified by the revenue variation. In other words, from the fact that the lobby maximizes at a different policy level, their bargaining power will be reduced at their revenue variation.

Under the consideration of the external pressures, we will observe the potential effects of combining both pressures under p_F as the set of final implemented policy in a legislature. In this case, we assume that both the politician receives an impact in the state-probability belief ($\pi_N \neq \pi_P$) and there is a lobby in the game. By comparing promised and final implemented polices:

$$p_0 - p_F = \bar{\gamma_N} - \frac{c\psi + nd\frac{(1-\beta)}{\beta}\bar{\gamma}_P}{c + nd\frac{(1-\beta)}{\beta}} = \frac{c(\bar{\gamma}_N - \psi) + nd\frac{(1-\beta)}{\beta}\left[\epsilon\left(\Delta_{\gamma}^{B,A}\right)\right]}{c + nd\frac{(1-\beta)}{\beta}} \tag{1.19}$$

The difference between the promised and the final implemented policies is reduced to the weighted average of two factors: distance between the social most preferred policy at population state-probability and maximizer revenue policy level, $\bar{\gamma}_N$ and ψ; and the variation of state probability belief in social most preferred policies between states, $\epsilon(\Delta_{\gamma}^{B,A})$. Each term is weighted by c and $nd\frac{(1-\beta)}{\beta}$, respectively.

Although individuals share the same utility functional form, they are heterogeneous with respect to the most preferred policy. In other words, any deviation in the final implemented policy compared to the promised one does not generate the same shock in all individuals. Consider an individual i that prefers most the policy level $\bar{\gamma}_i$. When the individual i observes the final implemented policy p_F, their expected utility function will change in relation to the promised policy p^0 as:

$$E[u_i(p_F)] - E[u_i(p_0)] = d\left[(\bar{\gamma}_i - p_0)^2 - (\bar{\gamma}_i - p_F)^2\right] \tag{1.20}$$

At p_F, an individual i will receive an impact in their expected utility level equivalent to $E[u_i(p_F)] - E[u_i(p_0)]$. In this case, if an individual will increase their expected individual utility at the final implemented policy, it will be since their most preferred policy is quadratically closer to the final implemented policy than the promised policy, $(\bar{\gamma}_i - p_F)^2 < (\bar{\gamma}_i - p_0)^2$. It supports that the shock of the final implemented policy in the individual utility function will depend on the location of the individual most preferred policy. Notice that the individuals with $\bar{\gamma}_i = \bar{\gamma}_N$ receive a negative impact in their utility function when the final implemented policy differs from the promised one, since it is the where the utility function is maximized.

From the previous section, individual trust in the national parliament was defined as a function that shapes:

$$\mathbb{T}_i(\bar{\gamma}_N) = Pr\Big(E[u_i(p_F)] \geq E[u_i(p_0)]\Big|\bar{\gamma}_N\Big)$$

In a few words, individual trust in national parliament is described as the likelihood of obtaining a higher individual utility with the final implemented policy than the promised one. From timing, the individuals are informed about their individuals' parameters ($\bar{\gamma}_i$), the social parameters ($\bar{\gamma}_N$, n and d), and the politician characteristics (β)[6], but they do not have precise information about ψ, c and ϵ, since it will be revealed in next stages. We denote $\mathbb{1}$, as the dummy variable that captures the situation where the expected utility level of the final implemented policy exceeds the obtained from the promised policy:

$$\mathbb{1} = \begin{cases} 1 & \text{if } E[u_i(p_F)] \geq E[u_i(p_0)] \\ 0 & \text{if } E[u_i(p_F)] < E[u_i(p_0)] \end{cases}$$

By considering the previous variable, the trust function for an individual i will be computed as:

$$\mathbb{T}_i(\bar{\gamma}_N) = \sum_{p_F \in P} \mathbb{1} Pr(p_F) = \sum_{\epsilon \in \mathcal{E}} \sum_{\psi \in P} \sum_{c \in R^+} \mathbb{1} Pr(\epsilon, \psi, c)$$

where $\mathcal{E} = [-\pi_N, 1 - \pi_N]$ is the domain of ϵ.

Trust is expressed as the probability of those policies p_F such that individual i is better-off. From Equation 1.16, the final implemented policy depends on the existence of *external* (ψ and c) or *internal* (ϵ) pressures. Given that individual trust depends on the final implemented policy affected by both pressures, the individual trust function will be described as the probability that ϵ, ψ and c weakly improve the individual utility compared to the promised policy. From this perspective, the individual trust function could be defined as a *Lebesgue* measure that computes the probability of improving the

---

[6] There is no difference in the politician characteristics β that generates distortions in the electoral competition framework. In other words, individuals will not consider β to cast their votes among parties.

expected utility from the individual expectations about both pressures.

1.6 Results

Once the individual trust function is defined, it will be exposed under the presence of *internal* and *external* pressure. With analytical purpose, our results will be complemented with a simulation that replicates the individual trust per different average individual ideal policy, $\bar{\gamma}$.

This exercise will be supported by a special case of $n = 701$ individuals with political preferences depending on the expected state. If the future state is A, their preference distribution will be uniformly distributed between 0 and $\frac{7}{10}$, $\gamma_A \sim Unif[0, \frac{7}{10}]$; whereas if the future state is B, their preference distribution will be uniformly distributed between $\frac{3}{10}$ and 1, $\gamma_B \sim Unif[\frac{3}{10}, 1]$. Population believes that the state A will occurs with probability $\pi = \frac{1}{2}$, whereas the probability of the state B will be $(1 - \pi) = \frac{1}{2}$. In aggregate, the expected political preference is distributed uniformly among $\frac{3}{20}$ and $\frac{17}{20}$ ($\bar{\gamma} \sim Unif[\frac{3}{20}, \frac{17}{20}]$). By considering that the individual preference order in the distribution is equal among states, then $\Delta_{\gamma}^{B,A} = \bar{\gamma}^B - \bar{\gamma}^A$ will be equal to $\frac{3}{10}$ for all agents. Furthermore, the individual discounted factor is standardized at $d = 1$. With respect to the pressures, we will assume state-probability distortion is distributed uniformly between $-\pi$ and $1-\pi$, $\epsilon \sim Unif[-\pi, 1-\pi]$; whereas there is a lobby, with a maximizer revenue policy uniformly distributed between 0 and 1, $\psi \sim Unif[0, 1]$, and has a variation revenue c allocated in an inverse exponential distribution of degree 1, $c \sim inv\ exp(1)$. A proper explanation about how the individual trust is computed in the simulation appears in Appendix as Individual Trust Simulation.

1.6.1 Internal Pressure

Consider that there is a distortion in the politician's state-probability belief. Under no presence of lobby, an individual i will determine their individual trust by determining:

$$\mathbb{T}_i(\bar{\gamma}_N) = Pr\left[(-\epsilon\Delta_{\gamma}^{B,A})\left[-\epsilon\Delta_{\gamma}^{B,A} + 2(\bar{\gamma}_N - \bar{\gamma}_i)\right] \leq 0\right] \tag{1.21}$$

The individual knows their own political preference, $\bar{\gamma}_i$; the most social preferred preference, $\bar{\gamma}_N$; and the difference between the most preferred policies among states, $\Delta_{\gamma}^{B,A}$; but their individual trust will depend on the expectation of state-probability distortion. Notice that the previous individual trust function depends on an inequality that is satisfied if the state probability distortion is allocated among $\epsilon \in \left[0, \frac{2(\bar{\gamma}_N - \bar{\gamma}_i)}{\Delta_{\bar{\gamma}}^{B,A}}\right]$ for $(\bar{\gamma}_N - \bar{\gamma}_i) \geq 0$; and $\epsilon \in \left[-\frac{2(\bar{\gamma}_N - \bar{\gamma}_i)}{\Delta_{\bar{\gamma}}^{B,A}}, 0\right]$ for $(\bar{\gamma}_N - \bar{\gamma}_i) \leq 0$. Then:

Definition 2 *A* ***favorable internal pressure*** *is a state probability distortion ϵ located in the interval:*

- $\epsilon \in \left[0, min\left\{\frac{2(\bar{\gamma}_N - \bar{\gamma}_i)}{\Delta_{\bar{\gamma}}^{B,A}}, 1 - \pi_N\right\}\right]$ *for individuals that satisfies* $(\bar{\gamma}_N - \bar{\gamma}_i) \geq 0$.
- $\epsilon \in \left[max\left\{\frac{2(\bar{\gamma}_N - \bar{\gamma}_i)}{\Delta_{\bar{\gamma}}^{B,A}}, -\pi_N\right\}, 0\right]$ *for individuals that satisfies* $(\bar{\gamma}_N - \bar{\gamma}_i) \leq 0$.

From the definition, a favorable internal pressure satisfies the policy set identified in the inequality of Equation 1.21 and considers the state-probability distortion domain $\mathcal{E} = [-\pi_N, 1 - \pi_N]$. In this scenario, a favorable internal pressure increases the individual expected utility level, since the state-probability distortion generates a political movement towards their most preferred policy level. From this perspective, the individual trust function could be reduced as follows:

$$\mathbb{T}_i(\bar{\gamma}_N) = \begin{cases} Pr\left(\epsilon \in \left[0, min\left\{\frac{2(\bar{\gamma}_N - \bar{\gamma}_i)}{\Delta_{\bar{\gamma}}^{B,A}}, 1 - \pi_N\right\}\right]\right) & \text{if } \bar{\gamma_N} > \bar{\gamma_i} \\ Pr\left(\epsilon = 0\right) & \text{if } \bar{\gamma_N} = \bar{\gamma_i} \\ Pr\left(\epsilon \in \left[max\left\{\frac{2(\bar{\gamma}_N - \bar{\gamma}_i)}{\Delta_{\bar{\gamma}}^{B,A}}, -\pi_N\right\}, 0\right]\right) & \text{if } \bar{\gamma_N} < \bar{\gamma_i} \end{cases}$$

Notice that the individual trust will depend on two different facts: the distance between individual and social most preferred policy $(\bar{\gamma}_N - \bar{\gamma}_i)$ and the impact of the states in the most preferred policy $\Delta_{\bar{\gamma}}^{B,A}$. First, individuals with their individual most preferred policy different from social one will increase their individual trust since a potential favorable state-probability distortion could induce the politician to implement a policy closer to their most preferred policy. Furthermore, the impact of the states in the political preference sizes the shock of state-probability distortion in the individual trust function. For example, when the difference is small, individual trust will increase since the most preferred policy will be invariant between states, allowing any movement of the final implemented policy towards such level to be favorable for the individual trust. Otherwise, when the impact of the states becomes larger, then the most preferred policy will depend on the states, being susceptible to the states under any variation in the final implemented policy.

Figure 1.1 contains the individual trust behavior depending on the average individual's most preferred policy. For those individuals who are far away from the social most preferred policy, their individual trust level equals 0.5, since it equals the probability

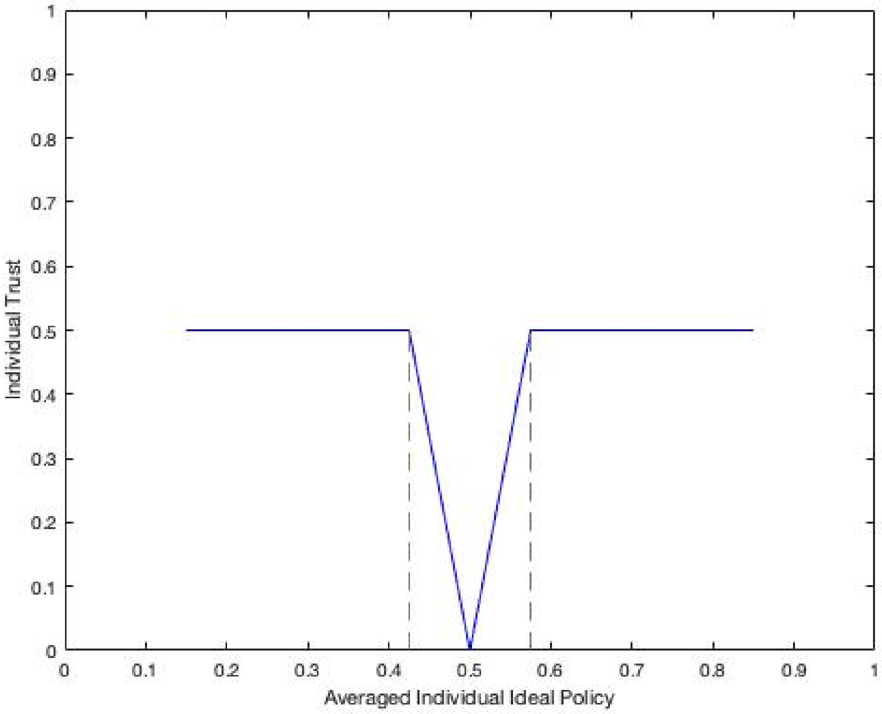


Figure 1.1: Effects of internal pressure on individual trust.

of obtaining favorable state-probability distortion. For example, those located at the left of the social most preferred policy need a negative state-probability distortion in order to increase their utility level. For those individuals whose most preferred policy is closer to the social most preferred policy, their individual trust level will depend on the probability of getting a favorable state-probability distortion. As an illustration, an individual whose average most preferred policy is located between both grey dashed lines will show a lower individual trust since a larger favorable state-probability distortion could reduce their individual expected utility level[7]. Last case, the individual located at the social most preferred policy reveals an individual trust level equal to 0, since the potential presence of a state-probability distortion will reduce their expected utility level.

1.6.2 External Pressure

Now, consider only the presence of the external pressure, excluding any effect coming from state-probability distortion. Notice that the population and the politician share the same state-probability belief; but, a lobby characterized by their maximizer revenue policy, ψ; and their variation revenue parameter, c, has incentives of suggesting a policy to the

---

[7] The location of the dashed lines is determined by observing when the improvement region exceeds the domain of state-probability distortion. In our case, given the distribution of the state-probability distortion, when an individual is characterized by either $-\pi < -\frac{2(\tilde{\gamma}_N - \tilde{\gamma}_i)}{\Delta_{\tilde{\gamma}}^{B,A}}$ or $-\frac{2(\tilde{\gamma}_N - \tilde{\gamma}_i)}{\Delta_{\tilde{\gamma}}^{B,A}} < 1 - \pi$, the required state-probability distortion will be any negative or positive value, respectively.

politician. From timing, individuals do not know precisely the lobby characteristics, but they can guess that if a lobby appears, the politician will accept their proposal. For this reason, individuals will expect some distribution with respect to the lobby characteristics. In this case, under no state-probability distortion, an individual i will determine their individual trust by:

$$\mathbb{T}_i(\bar{\gamma}_N) = Pr\left\{ \left(\frac{c(\psi - \bar{\gamma}_N)}{c + nd\frac{(1-\beta)}{\beta}} \right) \left[\left(\frac{c(\psi - \bar{\gamma}_N)}{c + nd\frac{(1-\beta)}{\beta}} \right) + 2(\bar{\gamma}_N - \bar{\gamma}_i) \right] \leq 0 \right\} \tag{1.22}$$

Due to timing, an individual knows their own political preference, $\bar{\gamma}_i$; their discount factor, d; population characteristics as the size of the population, n; and the most social preferred policy, $\bar{\gamma}_N$; and the politician factor, β, that measures the importance of politician about their own profit. With respect to the lobby characteristics, individuals are able to estimate them through their expectations. Given the uncertainty of two variables, we translate the individual trust as the probability that lobby revenue variation c, given their lobby maximizer profit policy ψ, increases the individual utility level. Notice that an individual will improve their utility if:

i) At any lobby revenue variation, when lobby maximizer profit policy is allocated between $\psi \in [\bar{\gamma}_N - 2(\bar{\gamma}_N - \bar{\gamma}_i), \bar{\gamma}_N]$ for $(\bar{\gamma}_N - \bar{\gamma}_i) \geq 0$ or $\psi \in [\bar{\gamma}_N, \bar{\gamma}_N - 2(\bar{\gamma}_N - \bar{\gamma}_i)]$ for $(\bar{\gamma}_N - \bar{\gamma}_i) \leq 0$.

ii) When lobby revenue variation does not exceed the threshold $\frac{2nd\frac{(1-\beta)}{\beta}(\gamma_N - \gamma_i)}{\bar{\gamma}_N - 2(\bar{\gamma}_N - \bar{\gamma}_i) - \psi}$ when lobby variation maximizer profit policy is allocated between $\psi \in [0, \bar{\gamma}_N - 2(\bar{\gamma}_N - \bar{\gamma}_i)]$ for $(\bar{\gamma}_N - \bar{\gamma}_i) \geq 0$ or $\psi \in [\bar{\gamma}_N - 2(\bar{\gamma}_N - \bar{\gamma}_i), 1]$ for $(\bar{\gamma}_N - \bar{\gamma}_i) \leq 0$. This situation will happen when $2(\bar{\gamma}_N - \bar{\gamma}_i)$ exists in the domain of ψ.

From both situations, we develop the following definition:

Definition 3 *A* ***favorable external pressure*** *is the lobby maximizer revenue policy ψ located in the interval:*

- $P_F^- = \left[max\left\{ \bar{\gamma}_N - 2(\bar{\gamma}_N - \bar{\gamma}_i), 0 \right\}, \bar{\gamma}_N \right]$ *for individuals that satisfies* $(\bar{\gamma}_N - \bar{\gamma}_i) \geq 0$.
- $P_F^+ = \left[\bar{\gamma}_N, min\left\{ \bar{\gamma}_N - 2(\bar{\gamma}_N - \bar{\gamma}_i), 1 \right\} \right]$ *for individuals that satisfies* $(\bar{\gamma}_N - \bar{\gamma}_i) \leq 0$.

Therefore, a favorable external pressure satisfies the policy set identified in the inequality of Equation 1.22 and considers the lobby maximizer revenue policy domain $P = [0, 1]$. If the individual is better-off at the lobby maximizer revenue policy level, any convex combination with the social most preferred preference will increase their expected utility level.

According to Definition 3, the individual trust is reduced to the following expression:

$$\mathbb{T}_i(\bar{\gamma}_N) = \begin{cases} Pr\left(\psi \in P_F^-\right) + Pr\left(c < c_{UB}|\psi\right) Pr\left(\psi \in [0, \bar{\gamma}_N - 2(\bar{\gamma}_N - \bar{\gamma}_i)]\right) & \text{if } \bar{\gamma}_N > \bar{\gamma}_i \\ Pr\left(\psi = \bar{\gamma}_N\right) + Pr(c = 0|\psi) Pr(\psi \in [0, \bar{\gamma}_N]) & \text{if } \bar{\gamma}_N = \bar{\gamma}_i \\ Pr\left(\psi \in P_F^+\right) + Pr\left(c < c_{UB}|\psi\right) Pr\left(\psi \in [\bar{\gamma}_N - 2(\bar{\gamma}_N - \bar{\gamma}_i), 1]\right) & \text{if } \bar{\gamma}_N < \bar{\gamma}_i \end{cases}$$

where $c_{UB} = \frac{2nd\frac{(1-\beta)}{\beta}(\bar{\gamma}_N - \bar{\gamma}_i)}{\bar{\gamma}_N - 2(\bar{\gamma}_N - \bar{\gamma}_i) - \psi}$.

Individual trust, under the unique existence of external pressure, is affected by many variables. First, population characteristics as the number of individuals, n; the policy discount factor, d; and the politician ratio, $\frac{(1-\beta)}{\beta}$. All of them have a positive effect on individual trust function since it increases the variation revenue threshold c_{UB}. This result comes from the fact that the three factors increase the transfers to induce the politician to implement the lobby suggestions. Notice that politicians with higher relevance in their own private gain (β closes to 1) induce to reduce the individual trust since it will ease any lobby suggestion at a cheaper transfer. Apart from previous factors, the individual's most preferred policy has an impact on the individual trust, depending on its position compared to the social most preferred policy. In this case, the larger the distance between both, the larger the effect on the individual trust function. In this case, individual preferences allocated in the extreme of the policy space will report a higher individual trust since the variation revenue threshold c_{UB} increases.

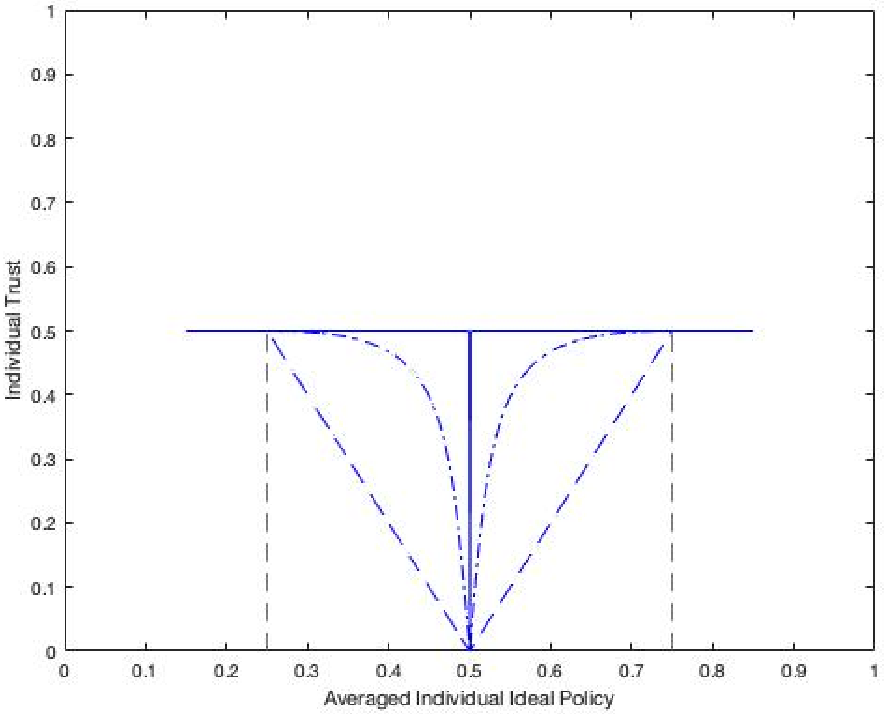


Figure 1.2: Effects of external pressure on individual trust.

Figure 1.2 contains the effect of external pressure on individual trust. There are three blue lines that characterize the effect of external pressure with different politician own-profit relevance parameter. For solid lines, the parameter is $\beta = 0.001$; for dashed-dots lines, $\beta = 0.991$; and for dashed lines, $\beta = 0.999$. Additionally, it is shown a grey dashed line that represents when the individual trust becomes constant.

There, it is estimated the effect of external pressure on individual trust levels at different politician own-profit relevance parameters. From low levels of this coefficient, $\beta = 0.001$, individual trust remains unchanged among the population. At this parameter level, the politician utility is mostly based on the social welfare level achieved by the finally implemented policy; making costly the transfer level that the lobby should provide to the politician to implement their suggestion. Given that there is a low likelihood of accepting a proposal, the individual whose most preferred policy is the social one will set their individual trust at 0. Increasing the politician own-profit relevance parameter at a higher ratio, as $\beta = 0.991$, reduces the individual trust below 0.5 for individuals where their most preferred policy is closed enough to the social most preferred policy. Notice, in this case, that the politician utility is mostly based on their own-profit level achieved by the finally implemented policy; allowing lobbies to persuade the politician to implement their suggested policy since optimal transfers are feasible. But, when the politician cares purely about their own-profit level, when $\beta = 0.999$, the individual trust will depend on the probability of identifying a favorable lobby variation revenue, c; and the maximizer profit policy level, ψ. In other words, attracting politicians to accept lobby proposals is costless, due to the propensity of the politician to consider more their own-profit level, inducing lobbies to suggest the final implemented policy. Notice that out of the set limited by the grey dashed lines, there are the individuals where individual trust achieves the maximum level. In this case, their individual trust level will only depend on the probability of finding a lobby with a maximizer profit policy level closer to the most preferred policy than the social one ($|\bar{\gamma}_i - \psi| < |\bar{\gamma}_i - \bar{\gamma}_N|$).

1.6.3 Internal and External Pressure

Now, consider the effect of both pressures on the behavior of the individual trust. In other words, the politician is able to perceive a different state-probability belief than the population one; and/or there could exist a lobby that suggests a different policy to be finally implemented. Then, under the existence of both pressures, an individual i will set their individual trust as follows:

$$\mathbb{T}_i(\bar{\gamma}_N) =$$

$$= Pr\left\{ \left(c(\psi - \bar{\gamma}_N) + nd\frac{(1-\beta)}{\beta}\left[\epsilon\left(\Delta_{\gamma}^{B,A}\right)\right]\right)\left[\left(\frac{c(\psi - \bar{\gamma}_N) + nd\frac{(1-\beta)}{\beta}\left[\epsilon\left(\Delta_{\gamma}^{B,A}\right)\right]}{c + nd\frac{(1-\beta)}{\beta}}\right) + 2(\bar{\gamma}_N - \bar{\gamma}_i)\right] \leq 0 \right\} \tag{1.23}$$

From timing, individuals know all variables except the state-probability distortion and lobby characteristics, as their revenue variation and maximizer profit policy level. Given the uncertainty of the three variables, their expectations will determine the individual trust level. Under this situation, an individual will improve their utility if:

i) At any lobby variation revenue, there are a favorable state-probability distortion and a favorable lobby maximizer profit policy. In this case, both variables will increase the individual utility level compared with the promised one.

ii) At a non-favorable lobby maximizer profit policy, the lobby variation revenue is smaller and there is a favorable state-probability distortion. Specifically:

   - When there is a non-favorable lobby maximizer profit policy on the left side. The lobby variation revenue must be non-negative and lower than $\frac{nd\frac{(1-\beta)}{\beta}\left(\epsilon\Delta_{\gamma}^{B,A}\right)}{(\psi - \bar{\gamma}_N)}$.
   - When there is a non-favorable lobby maximizer profit policy on the right side. The lobby variation revenue must be non-negative and lower than $\frac{\left[nd\frac{(1-\beta)}{\beta}\right]\left[2(\bar{\gamma}_N - \bar{\gamma}_i) + \epsilon\Delta_{\gamma}^{B,A}\right]}{\left[(\psi - \bar{\gamma}_N) + 2(\bar{\gamma}_N - \bar{\gamma}_i)\right]}$.

   Then, if the lobby variation revenue satisfies one of the previous cases, the finally implemented policy will improve the individual utility level.

iii) At any non-favorable state-probability distortion, the lobby variation revenue is greater and there is a favorable lobby maximizer profit policy. Specifically:

   - When there is a non-favorable state-probability distortion on the left side. The lobby variation revenue must be non-negative and higher than $\frac{nd\frac{(1-\beta)}{\beta}\left(\epsilon\Delta_{\gamma}^{B,A}\right)}{(\psi - \bar{\gamma}_N)}$.
   - When there is a non-favorable state-probability distortion on the right side. The lobby variation revenue must be non-negative and higher than $\frac{\left[nd\frac{(1-\beta)}{\beta}\right]\left[2(\bar{\gamma}_N - \bar{\gamma}_i) + \epsilon\Delta_{\gamma}^{B,A}\right]}{\left[(\psi - \bar{\gamma}_N) + 2(\bar{\gamma}_N - \bar{\gamma}_i)\right]}$

   Then, if the lobby variation revenue satisfies one of the previous cases, the finally implemented policy will improve the individual utility level.

iv) At any non-favorable state-probability distortion and lobby maximizer profit policy, when the lobby variation revenue is between $\frac{nd\frac{(1-\beta)}{\beta}\left(\epsilon\Delta_{\gamma}^{B,A}\right)}{(\psi - \bar{\gamma}_N)}$ and

$\frac{\left[nd\frac{(1-\beta)}{\beta}\right]\left[2(\bar{\gamma}_N-\bar{\gamma}_i)+\epsilon\Delta_{\gamma}^{B,A}\right]}{[(\psi-\bar{\gamma}_N)+2(\bar{\gamma}_N-\bar{\gamma}_i)]}$. Under this situation, the lobby variation revenue will improve the individual utility level.

From previous cases, the individual trust function as the probability that lobby variation revenue, c, improves the individual utility level with respect to the promised policy, conditional on the expectations of state-probability distortion ϵ and lobby maximizer profit policy ψ:

$$\mathbb{T}_i(\bar{\gamma}_N) = \sum_{\epsilon \in \mathcal{E}} \sum_{\psi \in P} Pr\left(c \in \mathbb{C}^+ | \epsilon, \psi\right) Pr(\epsilon) Pr(\psi)$$

where $\mathbb{C}$ is the domain of c that improves the individual utility level with respect to the one evaluated at the promised policy, specified previously (enumeration from i to iv).

The determination of individual trust function will depend on the lobby variation revenue characteristic that could offset the potential distortions in the state-probability and the lobby maximizer profit policy. Observing the main factors that determine the range of $\mathbb{C}$, we observe that the number of individuals, n; the discount policy factor, d; the politician ratio, $\frac{(1-\beta)}{\beta}$; and the difference between most preferred policies among states, $\Delta_{\gamma}^{B,A}$, have a positive effect on the lobby revenue variation threshold levels, increasing the influence of state-probability distortion on the individual trust. Then, any variation in these parameters could alter the effect of both pressures on individual trust. Furthermore, it is required to mention the effect of the distance between social and individual most preferred policy. For example, a reduction in the difference reduces the difference between both thresholds, decreasing the possibility of increasing the utility level when the state-probability distortion and the lobby maximizer profit policy are disadvantageous.

Figure 1.3 shows the effects both internal and external pressure on individual trust. There are three blue lines that characterize the effect of external pressure with different politician own-profit relevance parameter. For solid lines, the parameter is $\beta = 0.001$; for dashed-dots lines, $\beta = 0.991$; and for dashed lines, $\beta = 0.999$. Additionally, it is shown a grey dashed line that represents when the individual trust becomes constant.

There, it shows the individual trust under the presence of both pressures, under different politician own-profit relevance parameter. First, when this coefficient tends to 0 (e.g. $\beta = 0.001$), the individual trust will depend only on the effect of internal pressures, since convincing the politician is expensive. At this parameter level, the politician utility function is highly affected by the expected social welfare, increasing the level of transfers that politician would accept in order to implement the suggested lobby policy. This effect generates that individual trust function depends on the state-probability distortion, generating a curve similar to Figure 1.1. By increasing the politician own profit relevance,

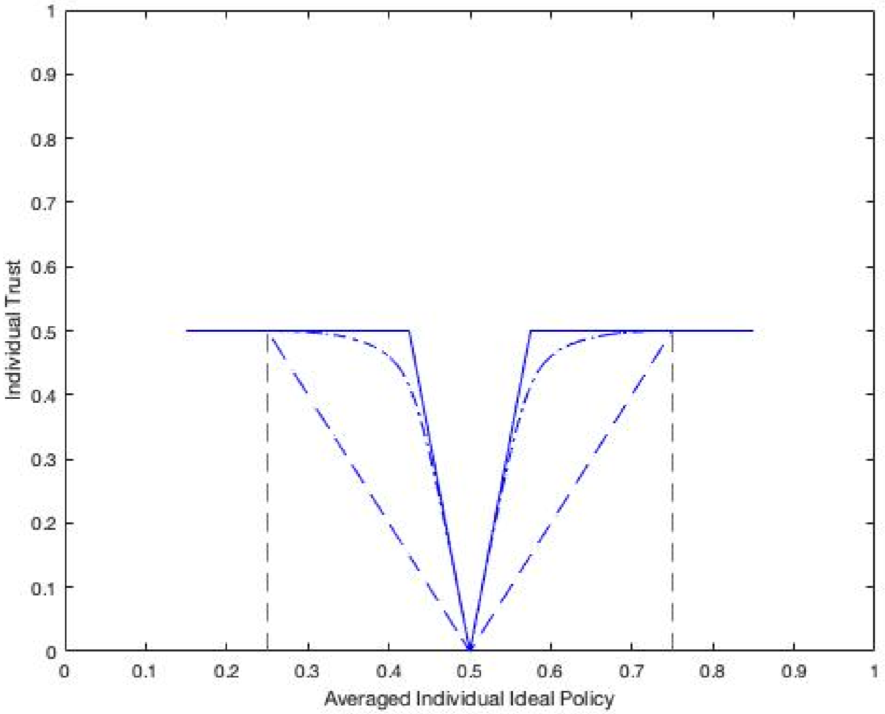


Figure 1.3: Effects of both internal and external pressure on individual trust.

the individual trust does not change until we achieve higher parameter level. Specifically, at $\beta = 0.991$, there is a reduction in the individual trust for those individuals closed to the averaged most preferred policy, $\bar{\gamma}_N = 0.5$, since cost of convincing politician to implement the lobby suggested policy reduces. The effect of politician own-profit relevance on the individual trust continues until we achieve the last result, where this parameter get closer to 1. There, $\beta = 0.999$, the politician utility function depends on their own-profit, being costless for the lobby to implement their suggestion as the final implemented policy.

In this section, the effect of pressures on individual trust is exposed. First, we started with internal pressure, where generates differences in the individual trust level among individuals whose most preferred policy is closer to the social one. Additionally, we have detected the factors that stress the variation of individual trust. Later, we have represented the effect of external pressure on individual trust is analyzed. There, the individual trust is affected by the expectations regarding the lobby maximizer revenue policy, ψ; and their variation revenue parameter, c. We observe that those politicians with a higher degree of own-profit relevance reduce the overall individual trust for the population where their most preferred policy is closer to the social one. Finally, when both pressures are combined, individuals are susceptible to more factors that affect their individual trust. Precisely, increasing the own-profit relevance parameter from the politician will "*erode*" the individual trust of the population whose most preferred policy is closer to the social one. In other words, this parameter will scale the effect of the internal pressure (against the external one). For example, for those politicians that care about social welfare, $\beta \rightarrow 0$, will increase the influence of internal pressures on individual trust. However, those politicians that rely on their own-profit $\beta \rightarrow 1$, will increase the effect of external effect on the individual trust.

1.7 Conclusion

This research provides a theoretical perspective on trust in national parliament, where individuals, given the winner policy level in electoral competition, identify the probability of obtaining, at least, an equal or higher outcome at the end of the legislature. With this purpose, we have defined the equilibrium in electoral competition and observed the potential effects of internal and external pressures in the final implemented policy. Starting from the internal pressures, the politician reveals a state-probability belief different from the population one. Under this situation, the politician will implement a set of policies that maximizes their own utility. With respect to the external pressures, the politician gets information about the lobby proposal, which it contains the proposed policy level and the utility transfer that will obtain if their suggestion is finally implemented. When a lobby proposes a policy, it will choose the level that maximizes its own profit function, considering the potential transfer. Last is necessary for the acceptance of the politician. According to these two pressures, the politician determines the final implemented policy that could differ from the promised one under electoral competition, affecting the individual trust function. Given the sequence of the game, and according to the individual most preferred policy, their level will depend on the expectation of the politician state-probability belief and lobbies characteristics.

From our model, we have observed that the final implemented policy equilibrium will be a convex combination between the most preferred policy at politician state-probability and lobby maximizer revenue policy level. Both will be averaged by the social and lobby characteristics. Whereas the former includes the number of individuals in the society, the policy discount parameter, and the politician ratio (representing the ratio between the relevance of expected welfare over the relevance of private gains in their utility level), the latter includes the lobby variation revenue parameter. In order to accept the offer, the lobby will compensate the politician by transferring the politician loss of accepting the offer.

After defining the trust in national parliament as the function that captures the likelihood of increasing the utility level in relation to the promised policy in electoral competition, we observed how it is affected by the presence of internal and external pressures. We concluded that the difference between most preferred policies among states could increase the inequality of individual trust through internal pressures, whereas the number of individuals, the policy discount parameter, and the politician ratio increase individual trust through external pressures. Additionally those politicians that mostly care about their private gains, $\beta \to 1$, reduce the individual trust for those individuals close to the median most preferred policy. When we combine both pressures, it magnifies previous results. As a result, individuals are subject to two pressures that will result in a reduced utility level.

Our main goal was to understand the potential factors that determine individual trust in national parliament. Thanks to the contribution of many different papers, we have been able to explain which factors could affect individual trust in national parliaments. We believed that this research could help to provide a theoretical explanation in many empirical pieces of research that aims to explain the trust in national parliament variations. Even, a potential future research is proving such trust in national parliament behavior in real data, as the one offered by the European Social Survey. If we consider that internal pressures are common among all countries, due to the strong networks in this globalized world; then the external pressure will result from a comparison analysis of trust in national parliaments.

As a policy implication, we have shown potential mechanisms to prevent deviations from promised policy. For example, the existence of a public lobby that protects this policy level through a higher enough variation in revenues, c, could balance the changes in state-probability belief or external pressures from lobbies with proposals different from the promised policy. In addition, as we have observed from many different pieces of research, the effect of media on trust in national parliament will exist through changes in population state-probability belief or their expectations concerning the existence of internal and external pressures. In other words, media is able to influence individual expectations to generate a future effect on the electoral turnout, through a reduction in the trust in national parliaments and increasing the cost of voting. And with this last effect in mind, trust in national parliaments has a potential effect in the electoral systems through its turnout. Due to this consequence, it is important to precise the potential conditions that affects any variation of such variable, as it is done in the current chapter.

Chapter 2

Too little clean air.

Existence of the Green Paradox under changes in the preference for energy.

(with Dr. Bipasa Datta)

2.1 Motivation

It is now widely acknowledged that the reduction of accumulated CO_2 from the process of obtaining and using fossil fuel has become one of the prominent features that countries have aimed to solve in the next years. According to conventional economic theory, there exists a wide range of potential policies that could reduce the incentive to generate such negative externalities, where most of them promote the internalization in the economic decision of agents (Coase, 1960). Despite the fact that these concepts are well understood and analyzed in the literature, Sinn (2008b) identified that imperfect policy designs worsen both environmental and welfare outcomes. Specifically, it is stated that: "*if suppliers feel threatened by a gradual greening of economic policies that would damage their future prices; they will extract their stocks more rapidly, thus accelerating global warming*". However, as mentioned in Di Maria et al. (2014), "*the literature virtually ignores the demand side of the resource market*". Indeed, they explained the difference between the use of the distinct types of energies attain to their marginal extraction cost or restrictions of their use, concluding that the demand-side does not affect the neglected consequences of an imperfect environmental policy.

Presumably, in accordance with the theoretical frameworks, there is coordination in imposing demand-side policies. As an example, policy-makers have begun to establish *Avoid-Shift-Improve* approaches to minimize the use of fossil fuel in favor of less polluted energies[1]. The intention of such type of policies is to induce the population to choose the alternative that reduces their impact on the environment and forms, consequently, stronger preferences for the use of clean energy. Apart from these incentives, it exists a huge academic contribution that shows that there is an increasing level of awareness of

---

[1] For more information, see Santos et al. (2021)

the use of renewable energy [2]. It is possible that the main factor is the political activity by promoting greener alternatives, but it could be also influenced by the justice considerations in a country (Tabi and Wüstenhagen, 2017); by the younger generations perspective about green technology (Tranter, 2011); by the more environmentally orientation of women than men (Ramstetter and Habersack, 2020); or by the political preference (Karlstrøm and Ryghaug, 2014). The influence of socio-demographic characteristics appears to be important in individual acceptance and, some of them, show that the social awareness of clean energy will rise in the future. Although policy-makers could have incentives of designing political strategies to prevent the accumulation of carbon dioxide, the major contribution of the social preference change could be attributed to the natural circumstances.

In the *Agenda 2030*, the *United Nations General Assembly* has committed to improving world welfare by achieving certain specific goals that promote peace and prosperity for people and the planet. Among their 17 *Sustainable Development Goals* (SDGs), the 13 SDG is attributed to these goals that mitigate the climate change impacts. Specifically, through the 13.2 SDG, it persuades the "*integration of climate change measures into the national policies, strategies, and planning*", intending to reduce the greenhouse gas emissions per year. Furthermore, the 13.3 SDG, it promotes education and awareness-raising on climate change mitigation, by stimulating countries to educate the population about sustainable development. In other words, *United Nations General Assembly* has aimed to internalize directly the environmental status of countries in the policy-maker decisions, promoting, at the same time, the increasing preference towards the use of renewable energy. Indeed, the international institution has included, as the 7 SDG, the increase of the affordable renewable share in the total final energy consumption. In principle, if policy-makers and the population consider the reduction of greenhouse gas emissions in their strategies or decision of consumption, respectively; then, affordable renewable energy will help them to achieve their desired goal. However, if we consider the supply perspective, there are a few questions that emerge from previous statements. For example, does the reduction in the cost of renewable energy halt the extraction of fossil fuel resources from the supply-side? Or, to what extent could an "*affordable*" clean energy prevent a rise in greenhouse gases? Nevertheless, the reduction in costs of a "*preferred*"

---

[2] There are many empirical analyses about the public support in renewable energy as Firestone et al. (2018) that shows that there is an acceptance in US regions for the wind projects; Tabi and Wüstenhagen (2017) where they analyze the factors that influences the social approval of hydro-power projects in Switzerland; Strazzera and Statzu (2017) showing the social acceptance in Photovoltaic and Building Integrated Photovoltaic technologies among different Mediterranean countries; and Sengers et al. (2010) studying the support for green energy technologies in Europe.

clean energy could stimulate the suppliers to reduce the price of fossil fuel resources, by revitalizing the extraction of such type of energy. Therefore, if the intention of the *United Nations General Assembly* is reducing the accumulation of CO_2 in the atmosphere, do their policies have a counter-factual effect and induce an augment of the emissions of greenhouse gases, provoking the existence of the so-called *Green Paradox*?

Our paper helps us to understand the emergence of the *Green paradox* when demand-side policies are established. This concept was defined by Sinn (2012) as the contradictory effect of green environmental policies that contribute indirectly to the acceleration of resource extraction. In specific, the contribution supports that, for the existence of the such phenomenon, the supply needs to react under the influence of a preferred imperfect substitute energy. The size of the previous response will depend on different conditions, such as, for example, the relative cost of renewable energy production and the market power of suppliers. In order to prove the previous contribution, we have set a two-period market model where both renewable and non-renewable energies are traded at each period of time. First, we consider that population (or consumers) is represented by their inclination to the use of renewable energy. And second, we assume that firms produce both types of energy and keep to maximize their two-period time profit function. After describing the concept of competitive equilibrium and identifying in our model, we stress the market outcome by stressing the consumer's preference for the use of renewable energy.

As a first result, we have detected that, if the economy is not able to exhaust all their resources, the exogenous shock generated in the consumer preferences contributes positively to environmental welfare, if and only if, it does not induce the incorporation of renewable energy as an alternative source. Otherwise, the presence of a substitute will stimulate the extraction of fossil fuels, producing a rise in the current level of emissions. Additionally, if the cost of the alternative is relatively cheaper, it will extend the increase of emissions for the next periods, being able to worsen their net present value.

As a second result, we have observed that, if the economy is able to exhaust all their fossil fuel resources, the exogenous shock generated in the consumer preferences will anticipate the extraction of fossil fuel to the current period. Even, the introduction of renewable alternatives will anticipate quickly the level of extraction. However, this result is subject to the extraction of fossil fuel in a discrete period of time, not allowing to postpone their decision extraction in even more future periods.

Apart of these results, we have detected that perfectly competitive markets are the ones that concentrates higher levels of extraction, when it is compared to monopoly.

This result holds also for current periods of those economies that fully exhaust their resources, observing a higher drop in future periods. However, under introduction of renewable energies, monopoly are susceptible to observe a higher increase in extraction, as an optimal price strategy. Additionally, we have observed the potential effect of these market structures in welfare.

A deep analysis of our previous results will be explained as follows. We will start with a description of the literature in section 2.2. Later, the framework will be introduced by detailing the different players and their respective roles in section 2.3. With an introduction of the potential effects of more green preferences in the amount of polluting energy in section 2.4, the competitive equilibrium will be identified in section 2.5, where different scenarios will be studied. First, we consider perfect competitive markets in subsection 2.5.1, whereas the market-power will be established, in form of monopoly, in subsection 2.5.2, with a detail comparison in subsection 2.5.3. Finally, section 2.6 describes, in detail, our results, providing some political implications and future research to consider.

2.2 Literature Review

Although it is complex to identify the origin of the enormous interest in the effectiveness of climate change policies, one could start with the discussion of Bohm (1993) with respect to the optimal alternative policy to deal with the phenomenon of *carbon leakage*. This concept was perfectly described by the model from Hoel (1994), where a group of countries cooperates to decide the optimal tax rate of production and use of fossil fuels. Depending on its size, a cooperative group may induce an increase in emissions from the non-cooperative countries (carbon leakage), since the reduction of the supply generated from the tax could rise their incentives to increase the current extraction. As a finding, Hoel (1994) showed that the sum of the optimal tax rate of production and use of fossil fuels were equivalent to their external marginal cost (*Pigovian* rate). Additionally, the optimal mix of supply and demand policies will depend on their respective elasticities. Following this model, Harstad (2012) shown that a potential first-best outcome to deal with climate change is the acquisition of the resources at the highest cost from the cooperative group of countries and keeping them permanently out of extraction.

However, if we return to Bohm (1993), there is explained the relevant role of supply-side policies to deal with climate change. This set of policies, aimed at reducing the production or extraction of fossil fuels (but not their use), results in a lower price sensitivity than, for example, reducing the demand for this resource. In this research, the authors have detected that a coordinated policy where the supply and demand of fossil fuels are reduced equally could prevent a fall in both market prices and producer profits. At the same period of time, Sinclair (1992, 1994) defined the concept of non-renewable as the type of resource where their fundamental decision is not how much to produce but when to extract it, as it occurs with fossil fuels. Apart from this concept, it was observed that if an ad-valorem non-renewable resource tax declines over time, owners or suppliers could postpone their extractions, reducing the potential current emissions. This result was subject to some assumptions as a specific production functional form where damages appear multiplying; pollution stock does not decay; and fossil fuels (as capital and oil) are substitute inputs in the Cobb-Douglas function. Thanks to Sinclair (1992, 1994), there started to appear many contributions around the concept of non-renewable resources [3], helping the literature to understand that, at each point in time, the optimal price of emissions should be similar to the present value of the marginal damages of future emissions (Hoel, 2011b, van der Ploeg and Withagen, 2012).

From these insights about the effectiveness of climate change policies, academics attain the potential externalities of controlling the economical behavior around fossil fuels. They recognize that some climate change policies, intended to mitigate carbon emissions, might actually induce profit-maximizing fossil fuels owners to precipitate the extraction of their resources. Due to the contradiction of policy-makers intentions, *Hans-Werner Sinn* assigned the name of "*Green Paradox*" in his academic book Sinn (2012)[4]. Under the design of climate change mitigation programs to prevent emissions, the political tentative of reducing demand would diminish the incentive to use fossil fuels since the social cost of polluting is considered in the benefit function of suppliers. But, when a dynamic framework is under review, in conjunction with their discounted value, the firms would be more partial to extract in current periods rather than future ones. By attaining in the emissions generated from the production and consumption of non-renewable resources, Sinn (2008b) argues that the overall level and the timing of greenhouse gas emissions do matter in evaluating the size of climatic change and their cost in the aggregate economy. Another aspect that must be considered when we determine the effectiveness of climate policy is the overall time path of extraction of fossil fuel sources. For instance, if policy-

---

[3] For example, Chakravorty et al. (2006) where their model deals with the ceiling with respect to the total amount of emissions; Chakravorty et al. (2008) where they focus on the order of extraction among non-renewable resources, detecting that their level of pollution must not be a determinant of such order; Ulph and Ulph (1994) who argues the findings identified by Sinclair (1992, 1994), showing an optimal different tax rate path under some conditions; or Tahvonen (1995) where it explains the differences of optimal pollution control depending on the global warming damage functional form.

[4] Important to consider his previous work in Sinn (2008a) and Sinn (2008b).

makers focus on reducing the demand for non-renewable resources, without interrupting the cumulative supply, it will lower the price for fossil fuels in the short run, increasing their emissions and non-affecting the cumulative level. Under the previous logic, Hoel and Kverndokk (1996) and Sinn (2012) state that to mitigate the emissions of non-renewable resources, a potential reduction of the discounted value of profits for resource owners in the short run would induce them to postpone the current extractions.

In one of their academic papers, Sinn (2008b) discusses the potential alternatives to deal with global warming when the fossil fuels supply does not react. As an alternative, the author analyses the demand reduction strategies, a type of policy that affects directly the demand for non-renewable resources that involves a decrease in the market price of fossil fuels. Another relevant aspect is the neoclassical optimal growth model with global warming externalities, where it does not include an alternative resource. Sinn (2008b) argues that those resources are represented intrinsically in the demand of fossil fuels, for the imperfect substitute ones; or generate a political opposition since some of them use many resources, for the perfect substitute ones. But, the supply-reduction strategy analyzed is the carbon tax, a policy politically infeasible. From this motivation, some academics, such as Tahvonen (1997) or van der Ploeg and Withagen (2012), start including the use of backstops as an alternative to fossil fuels, to analyze the demand-reduction strategies and potential policies that emulates the effects of first best policies[5].

The notion of the Green Paradox was extended in Gerlagh (2011), where the effect of low-carbon energy sources and their imperfect substitutes were introduced in the conventional models. Among the different scenarios that he idealized, the author identified two common trends: the *weak* and *strong* Green Paradox. The former occurs when a certain climate policy encourages fossil fuel owner firms to extract faster resources, raising, in this manner, the current emissions. The latter appears when the strategies of policy-makers rush the exploitation of the last unit of fossil fuel resource, increasing the net present value of cumulative damages from global warming. In order to illustrate them, Gerlagh (2011) shows that when there are the assumptions of marginal extraction costs are constant over time and independent of the resource stock, and the supply of the backstop energy is unlimited; a reduction of the price of a perfectly substitute clean technology will generate the existence of both the weak and the strong Green Paradox. But, under the linear demand assumption, when extraction costs linear increasing in cumulative supply are considered, only the weak Green Paradox appears. Nevertheless, when linear demand

---

[5] Where it imposes at each point of time a tax emission that equals the capitalized value of the stream of marginal damages of emissions. See Hoel (2011a) and van der Ploeg and Withagen (2012).

and constant extraction costs are supposed, if the minimal of the increasing marginal cost of the substitute is lower than the extraction cost, then both types of Green paradox do not arise.

In the last decade, many climate change policies were examined to comprehend their relationship with the existence of the Green Paradox. Most of the principal studies were focused on the supply-side, for several reasons. First, the decision of exploiting fossil fuel resources is directly linked with the optimal firms strategies. Another reason, as mentioned previously in Sinn (2008b), is the favorable shifts of the fossil fuels demand (intended to reduce the amount of resources) lead to a decrease in the market price, motivating a rapid extraction of resources from fossil fuel owners.

One of the topics that captured a big interest in the literature was the effects of fiscal policies on the fossil fuel market behavior. Hoel (2012) studies the sensitivity of emissions to the carbon tax level that depends on the behavior of extraction costs. As a finding, they identified that increasing extraction costs with a larger range of carbon tax growth could postpone emissions. Grafton et al. (2012) studies the effect of subsidizing alternative energy on the extraction path of non-renewable energy. They identified that smaller demand elasticity will reduce the increase in initial extraction, being smaller the change in the exhaustion time. If we focus on the effect of unilateral caps in the extraction of fossil fuels, Eichner and Pethig (2011) showed from their 2-period and 3-countries model that it could increase global emissions during the implementation period. The authors split the demand into two different parts, where one captures the amount of resources required to produce a certain amount of good (named as *intertemporal elasticity of substitution*), and the other the distribution among abating and non-abating regions (expressed as *elasticity of demand of the resource*). From this analysis, they conclude that a sufficiently lower intertemporal elasticity of substitution or a sufficiently higher demand elasticity for the resource is crucial for the existence of the Green Paradox.

Another aspect that might influence the appearance of the Green Paradox is the delayed time between the announcement and implementation of a climate change policy. Sinn (2008b) observed that the announcement of a rising in the carbon tax could induce fossil fuels owner to extract their resources quickly, anticipating future emissions and worsening the climate outcome in the short run. A theoretical explanation was developed by Di Maria et al. (2012), where it designs a simple model with a single resource and constant extraction cost where an announcement of a climate policy induces an increase in extraction between their announcement and final implementation, emerging a green paradox, as proved in Smulders et al. (2012) and Jus and Meier (2015). When they extend to multiple resources that differ in their pollution intensity, it tends to replace them with dirtier ones[6]. These consequences were corroborated by Di Maria et al. (2014), who designed the first empirical research for the existence of the Green Paradox. By deeply investigating the announcement of the cap on sulphur dioxide emissions that occurred in the introduction of the Acid Rain Program in the United States. From the time series used in the study, they show a decrease in the price of coal, but there was not much evidence to prove the consequent increase in the amount of coal used.

Apart from the properties of climate change policies, the intrinsic features of technology extraction and backstops play a role in the existence of the Green Paradox. Strand (2007) analyses the likelihood of creating an alternative to fossil fuels, concluding that technological agreement that makes fossil fuels unnecessary in the future increases current emissions.[7] In another study, van der Ploeg and Withagen (2012) identified the same results for stock-dependent marginal extraction cost for fossil fuels and constant marginal cost for the alternative energy, but they show that the Green Paradox occurs if the backstop is sufficiently expensive such that fossil fuels resource is exhausted.[8] For the improvement of the backstop technology, Hoel (2009) observe that if it produces at constant cost clean energy, which is a perfect substitute for fossil fuels, then it would have Green Paradox effects.[9] The shape of the marginal cost of alternative energy has been criticized in the literature. When all researchers have considered a constant marginal cost of alternative energy production, Chakravorty et al. (2011) have proved that the greater use of alternative energy will imply an increase in their unit cost of production. For example, the creation of biofuels requires different grades of land, a fact that implies an increase in their cost when the production of renewable energy rises. From this point of view, Van Long and Stähler (2014) uses both types of energy (renewable and non-renewable) at the same time. The authors detected that when there is a fall in the production cost of non-renewables could generate income effects leading to a fall in the interest rate and an increase in the current extraction rate.

The consequences of the Green Paradox are extended to the demand-side, where

---

[6] Gerlagh and Liski (2011) shows the effect of a potential substitute with a time-to-build delay. They have observed that fossil fuels producers alter the introduction of the renewable energy substitute by producing more fuel and reducing their price.

[7] Hoel (2011b) and Hoel (2012) also support the idea that environmental damage appears depending on the relationship between the evolution of the rise of the extraction costs when total extraction rise. Furthermore, they show a correlation with the time of the returns to investing in substitute energy.

[8] When the backstop becomes cheap, there will not be incentives to continue extracting energy from fossil fuel resources, reducing the effects of the Green Paradox.

[9] In conjunction with Hoel (2011a), we can understand that in a situation where fossil fuels are cheaper than their alternative and different countries establishes climate policies at different ambition levels, climate costs may increase as a result of improved technology of substitute resources and the absence of an efficient global climate arrangement.

potential changes in the uses or preferences of both types of energy. Chakravorty and Krulce (1994) and Chakravorty et al. (1997) help us to understand that the path of extraction of a certain type of fossil fuel could be altered with or without political intervention. Given the certain degree of substitution between the different types of fossil fuels, doing political decisions over certain non-renewable energy could affect the use of the others. With the interest of analyzing the consequences of subsidies, Grafton et al. (2012) also analyzes the effect of demand elasticity in the existence of the Green Paradox. In their results, they show that a lower demand elasticity could reduce the impact on the exhaustion date of the fossil fuel resource and diminish the increase in current emissions.

It is proved that the extended literature on the Green Paradox is based on the consequences generated by supply-side policies since their results in the market are quite ambiguous. But, from our sight, we believe that those strategies based on changes in the consumer pattern have not been exhaustively analyzed yet, remaining some questions for future research. For this reason, our model will be based on a two-period model where the amount of fossil fuel provided where consumers and firms, who supply both renewable and non-renewable energy, coexist in an economy. With the determination of the competitive equilibrium, we will observe the potential effects of changes in the consumer preferences in the potential decision of firms to extract fossil fuels.

2.3 Model

Consider a two-period economy with three different goods: *brown*, or polluting, energy (x_t) and *green*, or clean, energy (y_t); and the numeraire good (z_t). In period $t = \{1, 2\}$, the economy produces all two goods through the multi-supplier firms that simultaneously produce the pair of output x_t^s and y_t^s, where brown energy is produced through an input e_t of fossil fuel. As a strong assumption, suppose that a unit of brown energy is obtained by one unit of fossil fuel $x_t^s = e_t$. As a simplicity, understand that each unit of fossil extracted, generates one unit of emission. For the production of this type of energy, the economy is endowed with a stock of fossil fuel equivalent to r_0. Therefore, at any stage $\tau \in \{1, 2\}$, the available fossil fuel resources will be equivalent to $r_\tau = r_0 - \sum_{t=1}^{\tau} e_t$. At the end of the last period, the economy will have produced an aggregate amount of brown and green energy, $X = \sum_{t=1}^{2} x_t$ and $Y = \sum_{t=1}^{2} y_t$, respectively; where the first will determine the amount of fossil fuel kept in the ground, $r_0 \geq X$. According to this resource constraint, it will be said that the economy will *fully* exhaust all their available fossil fuel resources if previous constraint is satisfied with strict equality. However, if the resource expression hold only with an inequality, showing that all fossil fuel resources are not extracted, then the economy will have *partially* exhausted them. Apart of the behaviour of resources, there are two players that interacts in this economy: *multi-supplier firms* and *consumers*.

Suppose there exists a representative multi-supplier firm who is in charge of producing both types of energy with different cost structures. Starting from clean energy, where their total cost is equivalent to $k(y_t^s)$, it will be characterized by a positive marginal cost, $k'(y_t^s) \geq 0$, and increasing in the amount of output supplied by the firm j, $k''(y_t^s) > 0$[10]. Additionally, this kind of energy will not be restricted by any exhaustion constraint. Quite different for fossil fuels, characterized by a total cost equal to $c(r_t)x_t^s$, where their total cost per unit, $c(r_t)$, will be a decreasing function of the remaining fossil fuel on the soil $c'(r_t) \leq 0$. According to these characteristics, the representative firm will seek to maximise the following expression:

$$\begin{aligned} \max_{x_t^s, x_t^s} \quad & \Pi(x_t^s, y_t^s) = \textstyle\sum_{t=1}^{2} \Big\{ \left[p_{x,t} - c(r_t)\right] x_t^s + \left[p_{y,t} y_t^s - k(y_t^s)\right] \Big\} \\ \text{s.t.} \quad & r_0 \text{ is given,} \\ & \textstyle\sum_{t=1}^{2} x_t \leq r_0 \\ & x_t, y_t \leq 0 \quad \forall t \in \{1, 2\} \end{aligned} \tag{2.1}$$

In Equation 2.1, it includes the benefit that the representative multi-supplier firm obtains from trading *brown* and *green* energy in the economy, at prices $p_{x,t}$ and $p_{y,t}$, respectively. Additionally, we will assume no discount of second-period profits due to the absence of market rate of interest. As mentioned previously, multi-supplier firm need to satisfy the resource constraint.

On the other side of the market, there is a representative consumer in the economy that, each period, derives utility from the energy consumption and the numeraire good in a quasi-linear manner, $U_t(x_t, y_t, z_t|\theta) = u(x_t, y_t|\theta) + z_t$. Considering $u(x_t, y_t|\theta)$, it is strict concave in their arguments (where $u_x > 0$, $u_y > 0$, $u_{xx} < 0$ and $u_{yy} < 0$) and ruled by a parameter $\theta \in (\underline{\theta}, \overline{\theta})$ that captures the preference inclination of individuals towards fossil fuels and clean energy. With respect to the quasi-linearity of $U_t(x_t, y_t, z_t|\theta)$, it will allow to capture any income change only in the demand of the numeraire good [11]. Additionally:

Definition 4 *The representative consumer becomes "more (or less) green", if after*

---

[10] It is argued from van der Ploeg and Withagen (2012) that given the *McKinsey global green house abatement cost* curve, the shape of the production cost of clean energy must be convex. One of the potential explanations is the exponential increase in their equivalence to the cost of abatement for those resources that are less polluted as it could be the comparison between cars aerodynamics, wind energy, or solar photovoltaic panels.

[11] As a simplicity, consider that this numeraire good is supplied in the economy without any production and its price, or value to obtain it, is equal to 1.

perceiving a change in their utility function, from U to U', their evaluation for a certain bundle $\{x_0, y_0, z_0\}$ results in:

$$\frac{U_x(x_0,y_0,z_0|\theta)}{U_y(x_0,y_0,z_0|\theta)} > \frac{U_x(x_0,y_0,z_0|\theta')}{U_y(x_0,y_0,z_0|\theta')} \quad \left(or \quad \frac{U_x(x_0,y_0,z_0|\theta)}{U_y(x_0,y_0,z_0|\theta)} < \frac{U_x(x_0,y_0,z_0|\theta')}{U_y(x_0,y_0,z_0|\theta')} \right)$$

In a few words, the representative consumer becomes more green if after perceiving a change in their utility and evaluating the same bundle, then it reveals a greater change in the marginal utility from clean sources than fossil fuels. By attributing the changes of the utility function to the preference parameter θ:

Assumption 4 *The representative consumer is considered to be "more (or less) green" for higher (or lower) preference parameter θ.*

By considering the representative consumer intertemporal utility function, the consumer will seek to select the bundle of fossil fuels, renewable energy and numeraire good such that:

$$\begin{aligned} \max_{x_t,y_t,z_t} \quad & \mathcal{U}\left(x_t,y_t,z_t|\theta\right) = \textstyle\sum_{t=1}^{2} U_t(x_t,y_t,z_t|\theta) \\ \text{s.t.} \quad & \textstyle\sum_{t=1}^{2} p_{x,t}x_t + p_{y,t}y_t + z_t \leq \Pi^*(x_t^s, y_t^s) \\ & x_t, y_t, z_t \geq 0 \quad \forall t \in \{1,2\} \end{aligned} \tag{2.2}$$

When consumers face the energy market, they maximize their utility function, subject to their budget constraint $p_{x,t}x_t + p_{y,t}y_t + z_t \leq \Pi^*(x_t^s, y_t^s)$, where $\Pi^*(x_t^s, y_t^s)$ is the maximum profit of the multi-supplier firm. According to this set-up, consumers will set their optimal bundle equal to $U_x = p_x$ and $U_y = p_y$.[12]

In addition, the concept of competitive equilibrium will be required to understand the potential implications of exogenous variations in the preference parameter θ in the production of both types of energy:

Definition 5 *A competitive equilibrium consists of the consumer's decisions (x_1, x_2, y_1, y_2, z_1 and z_2), the multisupplier firm's decisions (x_1^s, x_2^s, y_1^s and y_2^s) and prices ($p_{x,1}$, $p_{x,2}$, $p_{y,1}$ and $p_{y,2}$), such that the following conditions hold:*

1. *Given their profit structure, the multisupplier firm's decision solve the multisupplier firm's profit maximization problem described in Equation 2.1.*

2. *Given their preferences, the consumer's decisions solve the household's intertemporal maximization problem described in Equation 2.2.*

3. *The prices $p_{x,1}$, $p_{x,2}$, $p_{y,1}$ and $p_{y,2}$ clear the markets.*

---

[12] This result is achieved by optimizing the individual utility function for the three different goods, normalizing the price of the numeraire one at 1.

In this model, it is said to be in a competitive equilibrium if both consumers and multisupplier firms optimize their own utility and profit functions, respectively; and the market is clear under the prices structures achieved between them. However, the understanding of the Green Paradox will depend on how the climate change externalities is include in the social welfare. Consider that the social welfare is as follows:

$$W(x_t, y_t, z_t) = \sum_{t=1}^{2} U_t\left(x_t, y_t, z_t|\theta\right) - c(r_t)x_t - k(y_t) - d\left(x_t\right)$$

where $U_t\left(x_t, y_t, z_t|\theta\right) - c(r_t)x_t - k(y_t)$ is the market welfare, whereas $d(x_t)$ is the environmental damage function composed by the aggregate fossil fuel as the argument x_t times the damage parameter d.

Before proceeding with the analysis of the impact of the preference parameter degree θ in the supply of fossil fuel energy in the market, it will be important to mention some interpretation about the time structure of the model. Our economy is restricted in two periods of time, where period 1 could be interpret as the time up to the medium term, from now on *current* time, and period 2 as the very long term, from now on *future* time. Notice that scientists recommends to avoid the medium term increases in the world mean temperature, below two degrees Celsius, to stabilise the world climate and prevent potential damages. Actually, the current aggregate medium term emissions exceeds the threshold level of emissions set by the scientific approach. It is necessary, as a climate stabilisation effort, to reduce the world emissions in current period. Apart of the efforts in supply strategies, it seems that reducing the incentives of using fossil fuel is one of the most prominent strategy from the demand-side. However, the potential reaction of multi-supplier firms under this change in the demand of brown energy could have contradictory effects.

In the following section, the equilibrium will be identified, assessing about the potential impact of improving the preference parameter degree in the supply of brown energy in the market. Additionally, it will extend the explanation by considering the price-taker, later analised in perfect competitive markets, and price-maker condition, later analised in monopoly, from the multi-supplier firms.

2.4 Influence of preference parameter in the extraction of fossil fuel

Consider the competitive equilibrium in the model mentioned previously obtained from the optimisation problems described in Equation 2.1 and 2.2, and the clearing market condition. Under any variation of the preference parameter, the utility function will change their shape because of Assumption 4, moving to another competitive equilibrium, composed by a different composition of brown and green energy, apart of the numeraire good. Although any increase of the preference parameter could motivate the consumers to switch from brown energy to green energy, the response of multi-supplier firms could be different under the presence of an alternative substitute. Suppose that there is a situation where an increase in the preference parameter induces an increase in the aggregate supply of fossil fuel, $\frac{\partial X}{\partial \theta} > 0$; emerging the aforementioned *Green Paradox*. Even, this concept could be segregated in two different as Gerlagh (2011). *One*, the *weak Green Paradox* exists if any variation of the preference parameter stimulate a rise in the emissions in current periods, motivated by a faster extraction of fossil fuel and, consequently, production of brown energy, $\frac{\partial x_1}{\partial \theta} > 0$. *Other*, the *strong Green Paradox* exists when the potential effect of the degree preference parameter affects the net present value of cumulative damages from global warming. Since there is no discount and the effect of fossil fuel supply in the environmental damage is linear, the strong Green Paradox will be reflected by observing the variation of the aggregate supply of fossil fuel, as mentioned previously.

Under the consumer's perspective, the change in the preference parameter will change their demand of the goods in the market. Specifically, consumers will identify their optimal bundle by reaching through brown and green energy an interior maximum: $\mathcal{U}_x = p_{x,t}$ and $\mathcal{U}_y = p_{y,t}$[13]. According to this condition, the demand of both types energy will depend on the prices of the market and the preference parameter in the following manner:

$$x_t = \mathbf{x}^d(\underbrace{p_{x,t}}_{(-)}, \underbrace{p_{y,t}}_{(+)}, \underbrace{\theta}_{(-)}) \qquad\qquad y_t = \mathbf{y}^d(\underbrace{p_{x,t}}_{(+)}, \underbrace{p_{y,t}}_{(-)}, \underbrace{\theta}_{(+)})$$

where $\mathbf{x}^d$ and $\mathbf{y}^d$ express the functional form where $p_{x,t}$, $p_{y,t}$ and θ are affecting the demand of brown and green energy, respectively. Notice that the price of the own good will affect

---

[13] This is the first order condition of the problem mentioned in previous section by setting the marginal utility equivalent to the prices for each type of energy. Notice that the numeraire good price is normalized at 1 and it will absorb the income effect.

negatively to its demand; whereas the price of the substitute good will affect positively. Additionally, increase of the preference parameter θ will affect positively to green energy, whereas negatively to brown ones.

From the other side of the market, multi-supplier firms will reach their interior maximum by setting the first order condition. In this paper, both *price-takers* and *price-makers* will be under review. According to these considerations, the first order condition will be reached when the marginal cost of each type of energy and the price (for price-takers) or the marginal revenue (for price-makers) are equal. Due to the linear quadratic shape of the utility function, the marginal revenue will depend positively on own prices; but different with respect to the preference parameter, being positive for brown energy and negative for green energy. By starting with green energy, the supply will be affected in the same manner for both types of price setters. First, the green energy supply for price-takers will be affected positively by their own prices, $y_t^{s,PT} = \mathbf{y}^{s,PT}(p_{y,t})$, where $\mathbf{y}^{s,PT}$ represents the inverse of the marginal cost function with $\frac{\partial y_t^{s,PT}}{\partial p_{y,t}} > 0$ because of the upward sloping of MC function. Whereas, in the case of price-makers, the green energy supply for price-takers will also be affected negatively by their preference degree, $y_t^{s,PM} = \mathbf{y}^{s,PM}(p_{y,t}, \theta)$, due to their influence in the marginal revenue: $\frac{\partial y_t^{s,PM}}{\partial p_{y,t}} > 0$ and $\frac{\partial y_t^{s,PM}}{\partial \theta} < 0$. By considering the market clearance, either using the price-taker or price-maker condition of multisupplier firms, both will drive us, by comparing with the demand of green energy, to the an equilibrium consumer price function for renewable energy, $p_{y,t}$. Additionally, both arguments $p_{y,t}$ and θ will affect positively.

Now, consider the fossil fuel demand and substitute the equilibrium consumer price function for green energy. There is a reduced form of fossil fuel demand, $x_t = \mathbf{x}^d(p_{x,t}, \theta)$, where $\frac{\partial \mathbf{x}^d(p_{x,t},\theta)}{\partial p_{x,t}} < 0$. Additionally, consider the choke price for the last unit of fossil fuel extracted and sold in the market: $\bar{p}_{x,t} = \mathcal{U}_x(0, y)$. At this point, because of the linear quadratic shape of the utility function, the effect of the preference parameter in the fossil fuel demand results in: $\frac{\partial \mathbf{x}^d(\bar{p}_{x,t},\theta)}{\partial \theta} < 0$. However, the behaviour of fossil fuel supply will depend on two different scenarios, depending on the status of their fossil fuel resources in the next two periods.

First, suppose that multi-suppliers are not able to spend all their resources in the next two periods: $r_0 > X$. In this case, the multi-supplier firm will set their price or marginal revenue (depending on their price setter condition) equal to the marginal cost of extracting one unit of fossil fuel, $c(r_t)$. Notice that depending on the price condition of multi-supplier firms, their supply will behave different. Similarly to the green energy supply, either using price-taker condition, resulting in $x_t^{s,PT} = \mathbf{x}^{s,PT}(p_{x,t})$; or price-maker condition, $x_t^{s,PM} = \mathbf{x}^{s,PM}(p_{x,t}, \theta)$, they will both react in the same direction under any variation of their arguments. For example, both $x_t^{s,PT}$ and $x_t^{s,PM}$ will react positively to the fossil fuel prices levels, $\frac{\partial \mathbf{x}^{s,PT}(p_{x,t})}{\partial p_{x,t}} > 0$ and $\frac{\partial \mathbf{x}^{s,PM}(p_{x,t},\theta)}{\partial p_{x,t}} > 0$. Additionally, the preference parameter will affect to the marginal revenue, affecting to the supply of price-makers in a positive sign, $\frac{\partial \mathbf{x}^{s,PM}(p_{x,t},\theta)}{\partial \theta} > 0$. However, even if the demand incentives the reduction of prices, multi-supplier firm will keep to diminish the cost of extraction by dropping the units of resources obtained from the ground, and holding their first order condition. Potentially, prices could increase under this situation, but the economy will observe a drop of fossil fuels in both periods. Notice that the preference parameter θ is able to determine the level of renewable resources provided into the market. For lower levels of θ, only fossil fuel will be supplied; whereas the simultaneous supply of both types of energy will occur under sufficient higher levels of θ. However, any variation of the preference parameter that implies the introduction of green energy in the market will motivate a rise in the extraction and production of brown energy. The main reasoning is the reduction in prices from brown energy to become competitive with respect to the new alternative energy. Notice that the substitution effect generated from the presence of such type of energy will alterate positively the amount of brown energy in the market.

As a second scenario, suppose that multi-suppliers are able to spend all their resources in the next two periods: $r_0 = X$. In this case, the multi-supplier firm will respect the Hotelling rule between both periods, respecting the full exhaustion of resources between both periods. Additionally, the economy will reach the choke price at the end of the second period. Since the increase of the preference parameter will affect negatively to the choke price, there will be a reduction in the units spend in the second period. However, it will anticipate the difference in fossil fuel extracted in the first period, holding in this case the Hotelling rule.

Notice that depending on the capacity of exhausting all their fossil fuel resources, the economy will face different situations, emerging to some paradoxical situations. As an illustration, assume an increase of the preference parameter in an economy that is not able to exhaust all their fossil fuel resources. Under this situation, the amount of fossil fuel traded into the market will be reduced in both periods. However, it does not always happen when the more green preferences of consumers induces the introduction of green energy alternatives. Since there is a new substitute in the market because of the consumer's preferences, the amount of fossil fuel in the market will increase in order to become competitive through prices. Due to the presence of the green alternative substitute, it will provoke a rise of the fossil fuel extracted in one or both periods, affecting to the aggregate level and emerging the green paradox. Per contra, suppose an augment of the preference parameter in an economy that now is able to exhaust all their fossil fuel resources because of their capacity. Under this situation, a potential increase in the preference parameter will rise the current extraction of fossil fuel by reducing the future one. Notice that the variation will be equivalent, not affecting to the aggregate amount of fossil fuel extracted. The boost in the first period will contribute to the existence of the weak Green Paradox.

Once an explanation about the potential effect of the preference parameter in the extraction of fossil fuel is specified, an illustration will be applied to observe how the different variables affect the evolution of fossil fuel in the model mentioned. In the next section, some functional forms will be determined to identify the competitive equilibrium in the market and, then, observing the potential effect under a change in the preference parameter degree.

2.5 Competitive equilibrium

In this section, we evaluate the equilibrium path under changes in the preference parameter different competitive market structures. In order to develop the analysis, the functional form of the consumer's utility, the per unit extraction cost, and the total cost of clean energy sources will be specified:

$$U(x_t, y_t, z_t|\theta) = (1-\theta)\left[ax_t - b\frac{x_t^2}{2}\right] + \theta\left[ay_t - b\frac{y_t^2}{2}\right] - x_t y_t + z_t \tag{2.3}$$

$$C(r_t) = c_0 - c_1 r_t \tag{2.4}$$

$$K(y_t) = k_0 y_t + k_1 \frac{y_t^2}{2} \tag{2.5}$$

where $\theta \in \left(\frac{1}{b}, 1-\frac{1}{b}\right)$[14]; a, c_0, c_1, k_0 and k_1 are positive; $b > 2$[15]; and $c_0 \geq c_1 R_0$.

Previous Equation 2.3, 2.4 and 2.5 respects the properties mentioned in section 2.3. Even, the preference parameter respects Assumption 4, allowing us to interpret any increase of the preference parameter degree as a "*more green*" movement in the preference utility of consumers.

Now, with the perspective of fully understanding the behavior of multi-suppliers firms under any action of the demand-side, the two different types of competition: perfect

---

[14] Required for the concavity of the utility function. By restricting the parameter in this interval, it will result in a positive definite Hessian matrix.

[15] It will allow having a wide range of preference parameter to analyze

competition (or price-takers) and monopoly (or price-makers) will be considered to observe the potential effect of the preference parameter in the amount of fossil fuel supplied and demanded in the economy.

2.5.1 Perfect competition

Consider that multi-supplier firms face a perfectly competitive market. In this case, they will keep to optimise Equation 2.1, with the condition that they are price-takers. From now on, the super-index “PC” will indicate the competitive equilibrium under perfectly competitive markets. As shown in Appendix: Chapter 2 - Competitive Equilibrium in Perfect Competitive Markets, the economy will face different situation depending on the characteristics of consumers and multi-supplier firms:

- **No production of energy**: It will occur under preference parameter degrees that satisfies: $\theta \in \left(\frac{a-c_0+c_1r_0}{a}, \frac{k_0}{a}\right)$. In this situation, the competitive equilibrium will be the no production of any type of energy since it will not result as economically beneficial for multi-supplier firms.

- **Only production of green energy**: It will occur when $(1-\theta)a - \frac{\theta a - k_0}{\theta b + k_1} < c_0 - c_1 r_0$ is satisfied. Under this situation, the production of brown energy is not beneficial, since the marginal cost of extracting one unit of fossil fuel exceeds the potential benefit obtained from the supply of brown energy. However, it will be beneficial for the multi-supplier firm to keep supplying green energy in the market at $y_t^{PC} = \frac{\theta a - k_0}{\theta b + k_1}$, $\forall t \in \{1, 2\}$.

- **Production of brown energy**: It will occur when it is beneficial extracting fossil fuel, despite the potential cost that could bring such action: $(1-\theta)a - \frac{\theta a - k_0}{\theta b + k_1} \geq c_0 - c_1 r_0$. However, the simultaneous production of both energies will depend on the characteristics of the economy:

  - **Partial exhaustion of fossil fuel resources**: It occurs when not extracting all their fossil fuels resources, $r_0 > X^{PC}$, is the competitive equilibrium of the economy. But, the simultaneity of producing both types of energy will depend on:

    - **Only production of brown energy**: It will occur if next two conditions are satisfied:

$$r_0 > \frac{2(1-\theta)b + c_1}{(1-\theta)b + c_1} \frac{(1-\theta)a - [c_0 - c_1 r_0]}{(1-\theta)b + c_1} \tag{2.6}$$

$$\frac{(1-\theta)a - [c_0 - c_1 r_0]}{(1-\theta)b + c_1} > \theta a - k_0 \tag{2.7}$$

First, it is not optimal for the multi-supplier firm to fully exhaust all their resources under Equation 2.6. Additionally, as second, it is not optimal for the multi-supplier firm to provide green energy in the market under Equation 2.7. In other words, price does not exceed the marginal cost of start producing green energy. Under this two different conditions, the multi-supplier firm will produce brown energy in current periods equal to $x_1^{PC} = \frac{(1-\theta)a-[c_0-c_1 r_0]}{(1-\theta)b+c_1}$, whereas in future periods, they will produce $x_2^{PC} = \frac{(1-\theta)a-\left[c_0-c_1\left(r_0-x_1^{PC}\right)\right]}{(1-\theta)b+c_1}$, reducing the amount of fossil fuel extracted by $\frac{(1-\theta)b}{(1-\theta)b+c_1}$ compared to the current period.

- **Simultaneous production of both types of energy**: It will occur if next condition is satisfied:

$$r_0 > \frac{[\theta b + k_1]\,[2(1-\theta)b + c_1] - 1}{[\theta b + k_1]\,[(1-\theta)b + c_1] - 1}\,\frac{[\theta b + k_1]\,[(1-\theta)a - [c_0 - c_1 r_0]] - [\theta a - k_0]}{[\theta b + k_1]\,[(1-\theta)b + c_1] - 1} \tag{2.8}$$

But Equation 2.7 is not held. In other words, first, it is not optimal for the multi-supplier firm to fully extract all their fossil fuel resources. However, the multi-supplier firm will provide green energy into the market because the price of this type of energy exceeds the marginal cost of start producing it. Under this situation, the market will reach the equilibrium by supplying the green energy at $y_t^{PC} = \frac{\theta a - x_t - k_0}{\theta b + k_1}$, $\forall t \in \{1, 2\}$. Moreover, the amount of brown energy in the market in the current period will be equivalent to:

$$x_1^{PC} = \frac{[\theta b + k_1]\Big[(1-\theta)a - [c_0 - c_1 r_0]\Big] - [\theta a - k_0]}{[\theta b + k_1]\,[(1-\theta)b + c_1] - 1}$$

Whereas in the future periods will be:

$$x_2^{PC} = \frac{[\theta b + k_1]\Big[(1-\theta)a - \left[c_0 + c_1\left(r_0 - x_1^{PC}\right)\right]\Big] - [\theta a - k_0]}{[\theta b + k_1]\,[(1-\theta)b + c_1] - 1}$$

Again, the amount of fossil fuel produced and sold into the market in future periods will be $\frac{[\theta b+k_1][(1-\theta)b]-1}{[\theta b+k_1][(1-\theta)b+c_1]-1}$ times lower than the ones in the current periods.

- **Full exhaustion of fossil fuel resources**: It occurs when extracting all their fossil fuels resources, $r_0 = X^{PC}$, is the competitive equilibrium of the economy. But, the simultaneity of producing both types of energy will depend on:

- $*$ **Only production of brown energy**: It will occur if Equation 2.6 is not satisfied and:

$$\frac{(1-\theta)b+c_1}{2(1-\theta)b+c_1}r_0 > \theta a - k_0 \tag{2.9}$$

First, it is optimal to extract all the fossil fuel resources. Notice that the firm cannot achieve the level requires if they do not consider the amount available of fossil fuel. In this case, the optimal extraction strategy will set by considering the fossil fuel resources constraint. Additionally, as a second condition, there is no incentive to produce green energy since the price for the first amount of fossil fuel extracted will not exceed their marginal cost, Equation 2.9. According to these two conditions, the competitive equilibrium under this situation will distribute the brown energy in $x_1^{PC} = \frac{(1-\theta)b+c_1}{2(1-\theta)b+c_1}r_0$ in the current period, and $x_2^{PC} = \frac{(1-\theta)b}{2(1-\theta)b+c_1}r_0$ in the future periods.

- $*$ **Simultaneous production of both types of energy**: It will occur if Equation 2.8 and 2.9 are not satisfied. First, multi-supplier firms will keep fully extracting all their available fossil fuel resources since it will be optimal. Additionally, as a second condition, the production of green energy will be beneficial since the potential price for the first unit of this type of energy exceeds their marginal cost. Under this situation, the market will reach the equilibrium by supplying the green energy at $y_t^{PC} = \frac{\theta a - x_t - k_0}{\theta b + k_1}$, $\forall t \in \{1, 2\}$. With respect to the amount of brown energy in the market, the one from the current period will be equivalent to:

$$x_1^{PC} = \frac{[(1-\theta)b(\theta b + k_1) - 1] + c_1(\theta b + k_1)}{2\left[(1-\theta)b(\theta b + k_1) - 1\right] + c_1(\theta b + k_1)}r_0$$

Whereas in the future periods will be:

$$x_2^{PC} = \frac{[(1-\theta)b(\theta b + k_1) - 1]}{2\left[(1-\theta)b(\theta b + k_1) - 1\right] + c_1(\theta b + k_1)}r_0$$

In all cases, under the competitive equilibrium, prices for brown and green energy will be equivalent to $p_{x,t}^{PC}(x_t^{PC}, y_t^{PC}) = (1-\theta)[a - bx_t^{PC}] - y_t^{PC}$ and $p_{y,t}^{PC}(x_t^{PC}, y_t^{PC}) = \theta[a - by_t^{PC}] - x_t^{PC}$. Once the price-taker behaviour of multi-supplier firms are exposed, it will be analysed the reaction of the competitive equilibrium in our model when multi-supplier firms act as price-maker entities. As an exercise to reduce the complexities of this type of market, as for example the strategic behaviour of competitors, the monopoly scenario will be considered in next subsection 2.5.2.

2.5.2 Monopoly

Consider that multi-supplier firm behaves as a monopoly in the market. As before, the super-index "M" will indicate the competitive equilibrium under monopoly markets. As shown in Appendix: Chapter 2 - Competitive Equilibrium in Monopoly, with the optimization problem described in Equation 2.1 and under the condition of price-maker, the economy will face different situation depending on the characteristics of consumers and multi-supplier firms:

- **No production of energy**: As in subsection 2.5.1, it will occur under a range of preference parameter degrees located in $\theta \in \left(\frac{a-c_0+c_1r_0}{a}, \frac{k_0}{a}\right)$. Remember that the competitive equilibrium will be defined as the no production of any type of energy under this situation, due to the lack of economic incentives to produce either one or both types of energies.

- **Only production of green energy**: This scenario will occur if $(1-\theta)a - 2\frac{\theta a-k_0}{2\theta b+k_1} < c_0 - c_1r_0$ is satisfied. Under this situation, no production of brown energy is optimal, since the marginal revenue of supplying one unit of this type of energy does not exceed the potential marginal cost of extracting it. With respect to the green energy, it will be produced at $y_t^M = \frac{\theta a-k_0}{2\theta b+k_1}$, $\forall t \in \{1, 2\}$.

- **Production of brown energy**: The presence of the brown energy in the market will occur if the potential marginal revenue obtained from supplying the first amount of brown energy exceeds the marginal extraction cost: $(1-\theta)a - 2\frac{\theta a-k_0}{2\theta b+k_1} \geq c_0 - c_1r_0$. But, under this situation, the potential supply of green energy and the level of resources will depend on the characteristics of the market:
  - **Partial exhaustion of fossil fuel resources**: This situation will occur if no extracting all their fossil fuels resources, $r_0 > X^M$, is the competitive equilibrium of the economy. However, the presence of the green energy will depend on:
    - **Only production of brown energy**: There are two conditions that need to be satisfied:

$$r_0 > \frac{4(1-\theta)b+c_1}{2(1-\theta)b+c_1}\frac{(1-\theta)a-[c_0-c_1r_0]}{2(1-\theta)b+c_1} \tag{2.10}$$

$$\frac{(1-\theta)a-[c_0-c_1r_0]}{2(1-\theta)b+c_1} > \frac{\theta a-k_0}{2} \tag{2.11}$$

From Equation 2.10, it indicates that it is not optimal to fully exhaust all the fossil fuel resources available. Additionally, from Equation 2.11, there is no incentive to produce green energy since the potential marginal revenue obtained from the production of this type of energy does not exceed its marginal cost. Under this competitive equilibrium, the amount of brown energy in the market will be equal to $x_1^M = \frac{(1-\theta)a-[c_0-c_1r_0]}{2(1-\theta)b+c_1}$, whereas in future periods, it will be $x_2^M = \frac{(1-\theta)a-[c_0-c_1(r_0-x_1^M)]}{2(1-\theta)b+c_1}$. Notice that the amount of fossil fuel will be $\frac{2(1-\theta)b}{2(1-\theta)b+c_1}$ times lower than in the current period.

∗ **Simultaneous production of both types of energy**: It precises to hold the following inequality:

$$r_0 > \frac{\Big[[4\theta b+k_1]\Big(2(1-\theta)b+c_1\Big)-8\Big]}{\Big[[2\theta b+k_1]\Big(2(1-\theta)b+c_1\Big)-4\Big]}\frac{[2\theta b+k_1]\Big[(1-\theta)a-[c_0-c_1r_0]\Big]-2[\theta a-k_0]}{\Big[[2\theta b+k_1]\Big(2(1-\theta)b+c_1\Big)-4\Big]} \tag{2.12}$$

However, Equation 2.11 is not satisfied. First, it is not optimal for the multi-supplier firm to fully extract all their available fossil fuel resources. Additionally, multi-supplier firms will have the incentive of providing green energy into the market since the marginal revenue of the first unit of this type of energy exceeds its marginal cost. At the competitive equilibrium, the amount of green energy will be $y_t^M = \frac{\theta a-2x_t-k_0}{2\theta b+k_1}$, $\forall t \in \{1,2\}$. With respect to the amount of brown energy, in the current period will be equivalent to:

$$x_1^M = \frac{[2\theta b+k_1]\Big[(1-\theta)a-[c_0-c_1r_0]\Big]-2[\theta a-k_0]}{\Big[[2\theta b+k_1]\Big(2(1-\theta)b+c_1\Big)-4\Big]}$$

Whereas in the future periods will be:

$$x_2^M = \frac{[2\theta b+k_1]\Big[(1-\theta)a-[c_0-c_1(r_0-x_1^M)]\Big]-2[\theta a-k_0]}{\Big[[2\theta b+k_1]\Big(2(1-\theta)b+c_1\Big)-4\Big]}$$

The amount of brown energy provided in future periods results in $\frac{[2\theta b+k_1](2(1-\theta)b)-4}{[2\theta b+k_1](2(1-\theta)b+c_1)-4}$ times lower than the ones in the current periods.

– **Full exhaustion of fossil fuel resources**: It happens when, under the competitive equilibrium, multi-supplier firms full extract all the available fossil fuel resources, generating $r_0 = X$. However, the types of energy provided in the economy will depend on the following:

- **Only production of brown energy**: This scenario will occur if Equation 2.10 is not satisfied and the following inequality is respected:

  $$\frac{2(1-\theta)b+c_1}{4(1-\theta)b+c_1}r_0 > \frac{\theta a - k_0}{2} \tag{2.13}$$

  From the first condition, multi-supplier firms will identify as optimal extracting all the fossil fuel resources. Notice that if the firm would not consider the resource constraint, they would have produced more than the available fossil fuel resources. In the second condition, Equation 2.13, multi-supplier firms will not provide green energy in the market, since the marginal revenue obtained from producing and supplying such type of energy will not exceed their marginal cost of production. Therefore, the competitive equilibrium will allocate the market with $x_1^M = \frac{2(1-\theta)b+c_1}{4(1-\theta)b+c_1}r_0$ units of brown energy in the current period, and $x_2^M = \frac{2(1-\theta)b}{4(1-\theta)b+c_1}r_0$ units of brown energy in the future periods.

- **Simultaneous production of both types of energy**: This last situation will result if Equation 2.12 and 2.13 are not satisfied. From the first condition, multi-supplier firm will exhaust all their resources as an optimal strategy. Whereas, as a second condition, it will be beneficial for the multi-supplier firm to produce green energy since the marginal revenue from the first unit will be weakly greater than their marginal cost of production. Therefore, the market will be in equilibrium by providing $y_t^M = \frac{\theta a - 2x_t - k_0}{2\theta b + k_1}$ units of green energy, $\forall t \in \{1,2\}$. Additionally, brown energy will be provided into the market in current periods at:

  $$x_1^M = \frac{2\Big[(1-\theta)b[2\theta b+k_1]-2\Big]+c_1[2\theta b+k_1]}{4\Big[(1-\theta)b[2\theta b+k_1]-2\Big]+c_1[2\theta b+k_1]}r_0$$

  And in future periods at:

  $$x_2^M = \frac{2\Big[(1-\theta)b[2\theta b+k_1]-2\Big]}{4\Big[(1-\theta)b[2\theta b+k_1]-2\Big]+c_1[2\theta b+k_1]}r_0$$

In all cases, under the competitive equilibrium, prices for brown and green energy will be equivalent to $p_{x,t}^M(x_t^M, y_t^M) = (1-\theta)[a - bx_t^M] - y_t^M$ and $p_{y,t}^M(x_t^M, y_t^M) =$

$\theta[a - by_t^M] - x_t^M$. After understanding the economic behaviour of the energy market, the preference parameter degree will be stressed to understand the environmental consequences generated.

2.5.3 Comparative analysis

Notice that depending on the characteristics of the market, multi-supplier firms could fully exhaust or not all their available fossil fuel resources. However, potential variations of the preference parameter could incentive the market to change their extraction strategies. But, these strategies could differ depending on the resource constraint. For analytical purposes, the status of the available fossil fuel resources will be under consideration to understand the reaction of the preference parameter in the extraction strategy of multi-supplier firms. First, consider that the characteristics of the market do not allow multi-supplier firms to extract all their fossil fuel resources:

Proposition 2 *In both types of market, when the fossil fuel resource constraint is not bound, any marginal increase in θ will induce a decrease of the amount of brown energy in equilibrium along periods, except if the marginal rise changes the preference parameter inducing the green energy in the market, then there will be an increase in the brown energy in the current periods. In addition, if the green energy introduced in the market is sufficiently cheaper (under lower levels of k_1), then there will be an increase in the amount of fossil fuel resources extracted.*

Proof: *See in Appendix, Proof 4.*

When the conditions in the economy do not allow the multi-supplier firm to exhaust all their fossil fuel reserves, any variation of the preference parameter degree from consumers will reduce the incentives of multi-supplier firms to produce brown energy. However, it will always occurs except when this green preference movement motivates the multi-supplier firms to introduce the green energy. Under the presence of a substitute good in the market, multi-supplier firms will increase the production of brown energy in the current periods. Since it also induces an increase in extraction and, at the same time, emissions, there will generate the existence of a weak Green Paradox. But, depending on the cost of the green energy, this reaction could be extended in future periods, augmenting the amount of fossil fuel extracted at the end of the second period. Due to the linearity in the environmental damage function and no discounting, it will generate an increase in the environmental damage, generating the strong Green Paradox.

Under this scenario, the condition of price-takers will allow the multi-supplier firms under perfect competition to generate more brown energy in their competitive equilibrium than the monopolist. However, it will generate a higher welfare level despite the potential environmental damage generated, compared, again, with the monopoly market. But, when the existence of the susceptibility of the strong Green Paradox is analysed, monopoly requires higher levels of k_1, compared with perfect competition, to stop increasing the brown energy production in future periods.

Now, consider that the characteristics of the economy allow multi-supplier firm to fully exhaust all their fossil fuel reserves available:

Proposition 3 *In both types of market, when the fossil fuel resource constraint is bound, any marginal increase in θ will increase the amount of brown energy in equilibrium in current periods.*

Proof: *See in Appendix, Proof 5.*

When multi-supplier firms exhaust all their resources, any increase in the preference parameter degree anticipates the extraction of fossil fuel. Due to the linearity in the environmental damage function, it will induce an increase in the emissions levels in current periods, emerging the weak Green Paradox. However, since fossil fuel resources are bounded by their availability constraint, it will reduce the amount of brown energy generated in the second period. Since there is no discount between periods and the environmental damage function is linear, then there will not be the presence of the strong Green Paradox. By comparing the amount of brown energy produced in both types of market under full exhaustion of resources, perfectly competitive markets are the ones that supplies more brown energy amount in the current period compared to monopoly. However, this last type of market will reduce the welfare in comparison to perfectly competitive markets.

2.6 Conclusion

This chapter focuses on understanding the implications of demand preferences in the use of energy. We characterized the equilibrium in a framework where two types of energy, brown (non-renewable) and green (renewable) sources, are supplied in the market. On the one side, we have parametrized the utility of consumers in a convex-linear function of the use of both energy sources. On the other side, we have defined the concept of multi-supplier firm as the responsible entity that provides energy in the market. After defining and identifying the equilibrium and analysing the potential effect of inducing more green in consumers, there are two results that are identified.

As a first result, if the characteristics of the economy partially exhaust all the fossil fuel resources, then more green preferences induce a reduction in the use of fossil fuel sources if and only if the changes in consumer preferences do not induce However, when more green consumers motivate the energy market to expand for renewable resources, the reaction in the use of fossil fuel energy is not clearly precise. Despite the inevitable negative response of multi-supplier firms in the quick extraction of fossil fuels in current periods after the introduction of the substitute alternative (*weak Green Paradox*), the prolonged variation could differ depending on some market factors. Specifically, for those cases where the new renewable resource is sufficiently cheaper, lower levels of k_1, it may induce an increase in the amount of fossil fuel resources extracted (*strong Green Paradox*). Additionally, it could motivate also an augment in the brown energy supply in future periods.

For a complete analysis, perfect competition and monopoly have been included to understand their potential effect in the environmental status. Although perfect competitive structures supply more amount of brown energy in the market, it restricts the existence of the strong Green Paradox. It contradicts the extremal imperfect market structures as monopoly is. In this case, as it is supported by the literature, it will help to reduce the amount of brown energy supplied in the market, but it will magnify their increase when clean energy is newly considered in the market. Since the multi-supplier firm influence the market outcome, it will result in as optimal strategy to reduce fossil fuel price even more, by exceeding the extraction of such resource.

As a second result, if the characteristics of the economy allow the full exhaustion of all the fossil fuel resources available in the beginning of the current period, then more green preferences induce a rise in the current brown energy in the market, in detriment of the amount produced in the future periods. However, this result is restricted to the two period times (current and future) used in our model. It is noticeable that the potential effect of increasing the preference parameter degree reduces even more the potential extraction levels in future periods, if no alternative is newly introduced in the market. But, since multi-supplier firms are restricted to spend all their full exhaustion resources, then it motivates the increase of brown energy in the second period. Depending on the increase of the preference parameter degree, it could allow the multi-supplier firm to stop fully exhausting their fossil fuel resources, moving to the first result detected in our model. With respect to the comparison between types of markets, it continues showing that perfect competitive markets exhaust faster the fossil fuel reserves than monopoly; however, it will contribute to a large welfare compared to the other type of market.

Our framework supports that a demand-side policy that interacts with consumer preferences could be imperfect if the goal is contributing to climate change. On the one side, since more green consumers motivate the introduction of renewable energy as an imperfect substitute for fossil fuel, it will provide incentives to increase the use of energy from this resource. On the other side, if multi-supplier firm fully exhaust all their fossil fuel reserves, more green preferences from consumer side will induce an anticipation of the brown energy supply in current periods. However, since it is imposed the fully exhaustion of fossil fuel, any variation in the preference parameter will allow a reduction of the extraction in future periods, forcing, at the same time, to a quick exhaustion of the fossil fuel reserves in current periods. In order to be precise with respect to the potential effect of the preference parameter degree in the extraction and emissions pattern of our model, the number of periods must be expanded, allowing multi-supplier firms to post-pone their extracting decisions in next periods.

In addition, our theoretical model provides some political intuition about how the policy must be implemented. For example, if motivating the use of clean energy is performed in conjunction with the contraction of their production cost or R+D policies, one of the consequences could be contradictory with the original goal of dealing with climate change. This idea contradicts one of the policies established by the *United Nations General Assembly*, where an increasing awareness of climate change, in *13.3 SDG*, is jointly promoted with the increase of *affordable* renewable energy in the total final energy consumption.

Apart from previous results, we leave the door open for many investigations related to this topic. As mentioned before, it is possible to clarify the results in a dynamic model with infinite periods of time. In this case, the economy will not be restricted to decide between exhausting or not their fossil fuel reserves in a the limited period of time, allowing them to act strategically according to their optimisation principles. Additionally, the introduction of imperfectly competitive markets, as oligopoly, could help us to increase our understanding with respect to the existence of the Green Paradox. Under this new approach, price-makers multi-supplier firms will internalise the strategic behaviour of competitors in their optimization problem. However, due to the complexity of these types of models, it would be important to analyse them under discrete time, as done in our model, and setting reasonably the anticipation of strategies from future periods.

Chapter 3

To phase out, or to phase down, that is the question. Stable coalitions under the Glasgow Climate Pact.

(with Dr. Bipasa Datta)

3.1 Motivation

In the month of November 2021, the city of Glasgow hosted the 2021 United Nations Climate Change Conference, also referred to as the 26^{th} Conference of Parties (COP26). Apart from holding the 16^{th} meeting of the Parties to the *Kyoto Protocol* (CMP 12) and the 3^{th} meeting of the Parties to the *Paris Agreement* (CMA 3), where adopts decisions and resolutions on the implementation of its provisions, it was also the first meeting, after the Paris Agreement, that enhance commitment towards mitigating climate change with new non-binding national plans, known as a national pledge, to achieve the climate-related targets for greenhouse gas emission reductions. After hosting the meeting for more days than expected, the 26^{th} Conference of Parties ends it up with the so-called *Glasgow Climate Pact*[1]. As a result, the agreement committed to halt and reverse forest loss and land degradation by 2030; reduce the methane emissions by 30 percent by 2030, compared to 2020 levels, through the Global Methane Pledge; and set out the determination for all new car and van sales to be zero-emission vehicles by 2040, globally and 2035 in leading markets, accelerating the decarbonization of road transport. Apart from such achievements, the *Glasgow Climate Pact* will historically be recognized as being the first arrangement to identify fossil fuels as one of the main drivers of global warming.

With the goal of reducing the potential consequences of climate change, all country members start the bargaining process to identify the best decision concerning the status of the unabated coal power. In an early stage, some countries were aimed to *Phase Out* the of coal, representing 46 percent of carbon dioxide emissions worldwide. However, some

---

[1] See Glasgow Climate Pact for more information.

nations, such as India or China, pledged to modify the agreement, reaching the following:

> "*Calls upon parties to accelerate the development, deployment and dissemination of technologies, and the adoption of policies, to transition towards low-emission energy systems, including by rapidly scaling up the deployment of clean power generation and energy efficiency measures, including accelerating efforts towards the phase-down of unabated coal power and phase-out of inefficient fossil fuel subsidies, while providing targeted support to the poorest and most vulnerable in line with national circumstances and recognizing the need for support towards a just transition;*"
>
> Glasgow Climate Pact, Article 36.

After lasting one day more than expected, the *Glasgow Climate Pact* agreed in encouraging countries members to *Phase Down* their unabated coal power, with a *Phase Out* of inefficient fossil fuel subsidies. It is generally perceived that this new specification is due to the lack of complete funding needed for developing countries to mitigate extreme weather events and build renewable energy infrastructure. Specifically, the COP26 welcomed around USD 350 million from the Adaptation Fund and around USD 600 million from the Least Developed Countries Fund. Nevertheless, developed countries came to the city of Glasgow, not achieving the promised USD 100 billion a year required from the developing countries. One of the perceptions of such international disagreement was the lack of financial aid. Some of the most important developing countries decided to reduce the restrictive measures proposed in COP26 to prevent any influence on their economy. Attaining this explanation, this paper will answer the question that under which conditions countries have incentives to *Phase Down* or *Out*. The decision around the unabated coal power resources will restrict countries by binding the potential amount of inputs that they will produce. Then, the potential effect of *Phase Down* or *Out* will affect the welfare optimization problem through their resource constraint. Additionally, this analysis will provide some intuition concerning the outcome obtained in the *Glasgow Climate Pact*.

With the previous goal, this paper will use some concepts from the International Agreement Environment literature. It will start by defining a model where countries can choose the level of fossil fuel (or dirty) and renewable (clean) energy. The countries, that will differ with respect to their preference parameter towards dirty energy, will have two different alternatives: acting as either an individual or cooperative entity. If they choose the first option, then they will maximize their own welfare function subject to the resource constraint. This optimization problem will follow two steps. First, countries decide the status of their resources and, second, choose the optimal amount of both types of energy to produce. However, if countries choose the second option, then they will be able to participate in an International Environmental Agreement that will contribute to them by inducing their pro-environmental behavior. When they join, all signatory countries will unanimously decide whether *Phase Down* or *Out* their dirty energy sources. As mentioned before, and a representation of the exogenous financial aid provided in the *Glasgow Climate Pact*, the countries will perceive an increase in the benefits obtained from such agreement through their pro-environmental behavior. Once decided the decision made concerning the resources, countries will individually choose their bundle of both energies to produce.

As a result of the model, countries will have the incentive to participate in the International Environmental Agreement if no imposition is made on the amount of resources that they could use. However, diminishing the level of dirty energy to use (*Phase Down*) will reduce the incentives for countries to participate in the International Environmental Agreement. Depending on the efforts made through the pro-environmental behavior, the *Phase Out* could be a potential equilibrium in the game, where the core is the stable coalition. Furthermore, it is discovered that the *Phase Down* is an equilibrium for lower pro-environmental behavior parameter. This paper will contribute to adding a new perspective of a new climate change bargaining process discussed in COP26, similarly as Murdoch and Sandler (1997a), Murdoch and Sandler (1997b) and Barrett (2006) in other Conference of Parties. Additionally, it will analyze a bargaining process where the signatory countries decide the upper-limit bound of their resource constraints, introducing the concepts of *Phase Down* and *Out* in the literature, and adapting concepts of stability in this bargaining process.

The potential explanation of such model will be as follows. After introducing the literature in section 3.2, a basic linear-quadratic model of an emissions game will be provided in section 3.3. First, the non-cooperative equilibrium will be identified in section 3.4; and, then, the cooperative equilibrium will be assessed in section 3.5, evaluating them in three different scenarios: no intervention, *Phase Down* and *Phase Out*. Once the results are presented, a final explanation, jointly with the political implications will be detailed in section 3.6.

3.2 Literature Review

The International Environmental Agreement has become a relevant topic due to the urgency of identifying a potential solution for the impact of climate change. Many papers have focused on the most important agreements in the last years. For example, Murdoch and Sandler (1997a) analyzed how the reduction of the sulphur emissions that emerged in the Helsinki Protocol was because of the non-cooperative behavior of countries. In others, as Murdoch and Sandler (1997b), showed the low influence of the Montreal Protocol in the chlorofluorocarbon emissions in the late 1980s. As it was proved, the emissions cutbacks were, in large part, explained by the national income and political and geographical factors. It is well worth mentioning the contribution of Barrett (2006) in the *Kyoto Protocol*, analyzing the lack of a self-enforcing agreement and offering a potential solution to enlarge the size of the coalition by promoting cooperative R&D or encouraging collective adoption of breakthrough technology.

With these papers in mind, the main goal of the current research is interpreting the outcome obtained in the last *Glasgow Climate Pact*. Due to the active research in the International Environmental Agreement, there is an extended literature captured in Finus and Caparros (2015) and Caparros (2016). As well mentioned in the last survey, the understanding of the International Environmental Agreement encompasses many different perspectives, as the stability of the agreements (Finus, 2003); the composition of the core in a burden-sharing rule framework (Helm, 2000); the dynamic behavior under climate change problems (Calvo Ramón and Rubio, 2013); or the outcome obtained for the negotiation process dealt (Caparros, 2016). Since our research is focused on the static approach of the *Glasgow Climate Pact*, it is necessary to comprehend the stability condition and the potential formation of the core.

Through the understanding of the cartel formation in a price-leadership model, D'Aspremont et al. (1983) split the concept of stability into two different categories: *internal* refers to the incentives of a member to leave the coalition; whereas *external* refers to the incentives of a non-member entity to join the coalition. As a result, the determination of stability will depend on the sensibility of profits under the addition or loss of a member would generate. Many different models combined this concept of stability in the International Environmental Agreement network. Carraro and Siniscalco (1993) identified that the instruments to implement cooperation matters. For example, self-financed transfers could induce an increase in the number of members. Even, they show that those efficient and effective instruments that discourage free-riding promote stability among countries. Barrett (1994) found, in a model with quadratic costs and benefits, that under the presence of an agreement, the level of cooperation will be higher when the gains of cooperation are smaller.[2] On the contrary, if the gains are large, the coalition will not sustain large cooperation. According to these ideas, Nordhaus (2015) show that adding excludability in the International Environmental Agreement could be more attractive. Under this perspective, it emerges the *small coalition paradox*, since countries will prefer a regime with penalties and modest carbon prices, rather than one with no sanctions. This result is shown in Hagen and Schneider (2021), where the introduction of import tariffs combined with retaliation by outsiders can stabilize large coalitions but also results in the non-cooperative equilibrium by destabilizing small coalitions. Additionally, the retaliation measures compromise the appeal of trade sanctions to enlarge coalitions.

There are many different ways of commitment in the International Environmental Agreement to reduce greenhouse gas emissions. From previously mentioned papers, as Carraro and Siniscalco (1993) or Barrett (1994), the signatory countries were restricted to compensating the loss generated from abatement with profit transfers. Apart from these contributions, Carraro and Siniscalco (1997) focus their analysis on the investment of cleaner technologies. Such type of investment is considered a club good, in the sense that only the signatory countries are able to perceive the technology spillovers. In Barrett (2006), it doubts the effectiveness of treaties focused on abatement, as done in the *Kyoto Protocol*. On the contrary, it stimulates the main focus on green technology rather than emission reduction. Specifically, it analyses the use of *breakthrough technology*[3], identifying that if such type of technology exhibits increasing returns, it could sustain the full participation of countries.

However, due to the motivation of our paper, the club good offered in the *Glasgow Climate Pact* was not mainly based on investing in R&D to reduce the potential consequences of greenhouse gas emissions. From our perspective, the *Glasgow Climate Pact* incentives the countries to join the coalition to reduce the potential extraction of unabated coal power with the use of green technology, because of their positive impact of contributing to the worldwide welfare. This concept is closely related to the model of giving that includes warm glow, modeled by Andreoni (1989, 1990). In these papers, individuals were assumed to contribute to a public good due to two reasons. First, there is altruism from individuals, in the sense that individuals demand more of the public good.

---

[2] Other papers, as De Zeeuw (2008) discusses the effect of the stable coalition in a dynamic approach. As a result, large coalitions are sustained if damage costs are very small compared with abatement costs.

[3] See Hoffert et al. (2002), it could be defined as the type of technology that offers the possibility of eliminating greenhouse gas emissions, as it is defined in Rubio (2017).

And, second, people perceive some private benefits from the public good. According to Andreoni (1989), even because of the second characteristic, the model of giving is also called impure altruism. Kahneman and Knetsch (1992) realized that the moral satisfaction of contributing to public goods is represented by the willingness to pay asked in contingent valuation. In the International Environmental Agreement, Chambers and Jensen (2002) designed a signaling game of donations between two countries, where one country perceives a warm glow by the fact of donating. After identifying the pooling and separating equilibria, they show, under certain conditions, that a country could misrepresent its type in order to receive more aid. Other studies have improved and related the warm-glow concept to environmental factors. According to Taufik and Venhoeven (2018), they represent the pro-environmental behavior as an impure altruist concept, since their eudaimonic view[4] and the moral behavior. While the former refers to the individual effect of the pro-environmental attitude, the latter considers the benefit in the nature and well-being of other people. By considering this concept, the *Glasgow Climate Pact*will be considered as a potential coalition that promotes the more use of clean energy for private gains of the public good, and its contribution to the reduction of the emissions that affects the rest of countries.

Thanks to the literature, our analysis will focus on the basic linear-quadratic model of an emissions game, as used in Rubio (2017). Apart from following the concept of stability, the use of the *Khun-Tucker* conditions used in the previous paper will help us to develop our paper by ensuring non-negative pay-offs. Additionally, we consider the presence of the pro-environmental parameter, following the concept of *warm-glow* defined in Andreoni (1989, 1990), for the signatory countries that join the International Environmental Agreement. Since the *Glasgow Climate Pact* influences the resource constraint of countries, then each of them will optimize their own welfare function subject to the imposed resource constraint. This will be their non-cooperative decision; however, the countries will have the opportunity of joining the International Environmental Agreement that benefits them through the use of clean energy, but will cooperatively restrict their resources under a coalition decision.

---

[4] Defined in Taufik and Venhoeven (2018) as "*the view that positive and negative emotions related to environmental behavior have their roots in the behavior being a meaningful experience*".

3.3 Model

Consider a static model with N countries or regions that produce two types of energy: *dirty* or *non-renewable*, x_i, and *clean* or *renewable* ones, y_i. Additionally, regions contribute negatively to the atmosphere by the production of dirty energy. The production and consumption of both types of energy will give rise the regional benefits, in a linear-quadratic manner:

$$B_i(x_i, y_i) = a_i^x x_i - \frac{x_i^2}{2} + a^y y_i - \frac{y_i^2}{2} - x_i y_i$$

where $a^y > d\delta$ and $a_i^x \in [d\delta, a^y]$[5]; being d the constant marginal climate damage, and δ is the proportion of emissions per extracted unit of dirty energy. For future simulations, the countries will be distributed as $a_i^x \sim Unif[d\delta, a^y]$.

The dirty energy is obtained from an exhaustible resource or reserves R, common among all regions, where the amount of input extracted (z_i) needs to be transformed in a certain way to obtain the final output (x_i) in the following manner: $x_i = (z_i)^\phi$, where $\phi \in (0,1)$. According to their reserves level, countries set the amount of resources that they can exploit by determining the percentage of available resources to extract, $\theta_i \in [0,1]$. Notice that the previous parameter determines the level of inputs extracted by a country i: $z_i = \theta_i R$. Depending on the level of the percentage of available resources to extract, there is a *Phase Out* scenario if $\theta = 0$; *Phase Down*, if $\theta \in (0,1)$; or *no intervention*, if $\theta = 1$. Furthermore, in order to produce a certain amount of the final output of dirty energy, the production cost, $C(x_i)$, is equivalent to:

$$C(x_i) = c\frac{x_i^2}{2}$$

where $c > 0$ represents the coefficient of the increasing marginal cost of extracting one unit of x_i.

Assume that each unit of dirty energy consumed in the country contributes negatively to the atmosphere in $e_i = \delta x_i$, where $\delta > 0$. The aggregate amount of emissions generated from this type of energy, $E = \sum_{i=1}^N e_i$, will reduce the environmental welfare at $D(E) = dE$, where $d > 0$ represents the constant marginal climate damage. As mentioned in Golosov et al. (2014), consider that linearity in the damage function is not too extreme a simplification, since "*the composition of a concave S-to-temperature mapping with a convex temperature-to-damage function may be close to linear*".

In addition to x_i, all regions are able to get energy from a non-exhaustible or clean

---

[5] These conditions are required to stimulate the consumption of dirty energy.

source y_i, that generates a quadratic production cost specified as:

$$K(y_i) = k\frac{y_i^2}{2}$$

where $k > 0$ represents the coefficient of the private marginal cost of producing one unit of y_i.

According to the previous characteristics of the game, the countries need to decide whether to cooperate or not in order to maximize their welfare function subject to the reserves restrictions by selecting the level of dirty and clean energy to produce. Depending on the type of cooperation that a country i chooses, it will face a certain timing.

Suppose that a country acts in a non-cooperative manner. Denote the optimal pair of both types of energy produced and consumed by $\{x_i^{NC}, y_i^{NC}\}$, being the pair of both types of energy that maximizes the welfare function, W_i^{NC}, subject to the resource constraint, $x_i = f(\theta_i R)$:

$$W_i^{NC} = B_i(x_i, y_i) - C(x_i) - K(y_i) - D(E_N)$$

As an alternative, all regions have the opportunity to join an International Environmental Agreement, which involves a cooperative game of three stages.

First, all regions decide whether to join or not in the International Environmental Agreement. If they decide to join, they become a signatory country and form part of the coalition M, being characterized by $j = 1, ..., M$. Otherwise, when they do not join the coalition, they act as a non-cooperative entity. The International Environmental Agreement allows countries to perceive the cooperation as a positive benefit represented by the monotonic transformation of the aggregate amount of renewable energy, $\beta \sum_{j=1}^{M} y_j$, where $\beta > 0$ and y_j is the clean energy generated by a signatory country j once it joins the cooperation. There are two different ways to understand the positive benefit generated by cooperation. First, as a purely altruistic warm glow behavior mentioned by Andreoni (1989, 1990), where the country gets the reward of their private decision through the global impact generated by the production of more renewable energy from the coalition. And, second, the pro-environmental behavior, detailed in Taufik and Venhoeven (2018), obtained from the increasing benefit of clean energy. When countries join the International Environmental Agreement, apart from perceiving the direct benefit of producing clean energy, they also influence the benefit of the rest of the countries by either reducing damage or increasing the value of clean energy produced in the coalition through the pro-environmental parameter multiplier.

Once all countries have determined their decision, all the signatory regions, cooperatively, decide the maximum percentage of dirty energy resource use, $\bar{\theta}^{IEA}$, that minimizes the worldwide aggregate damage level. The *Glasgow Climate Pact* attains to reduce the causes of climate warming; by respecting, promoting, and considering their respective obligations. Then, such minimization problem is interpretable according to the *Glasgow Climate Pact*, if all countries respect their own participation constraints.

Finally, as the last stage, the signatory regions will decide individually the pair of energy, $\{x_j^{IEA}, y_j^{IEA}\}$, that they finally produce and consume. Notice that all these decisions will be set by the welfare function of the signatory region that decides to be part of an International Environmental Agreement and will perceive a net benefit of energy consumption W_i^{IEA} equivalent to:

$$W_j^{IEA} = B_j(x_j, y_j) + \beta \sum_{j=1}^{M} y_j - C(x_j) - K(y_j) - D(E_N)$$

In other words, the International Environmental Agreement encompasses three stages that are specified in the following scheme:

Stage 1 Country i decides whether to join the International Environmental Agreement or not, becoming a signatory country in case of acceptance.

Stage 2 All signatory countries M decide, cooperatively, the maximum percentage of dirty energy resources to exploit, $\bar{\theta}^{IEA}$, that minimizes the worldwide aggregate damage level.

Stage 3 The signatory country $j \in M$ decides the optimal pair of both types of energy, $\{x_j^{IEA}, y_j^{IEA}\}$.

Given the different alternatives that regions have under their control, it will be required to analyze the stability condition:

Definition 6 *It is said that an agreement consisting of M signatories regions is stable if:*

- $W_j^{IEA}(M) \geq W_j^{NC}(M \backslash \{j\}) \quad \forall j \in 1, ..., M.$
- $W_i^{NC}(M) \geq W_i^{IEA}(M \cup \{i\}) \quad \forall i \in 1, ..., N - M.$

Following a similar definition as Rubio (2017), it is said that an agreement is stable if both conditions are satisfied. The first corresponds to the concept of *internal stability* condition, where all members of the agreement are, at least, better-off cooperating than acting as a single region. The second corresponds to the *external stability* condition, where all non-signatories weakly prefer their individual energy composition rather than being part of the cooperation. Due to the static construction of the game, the stability conditions will be focused on the internal ones.

In the next sections, the optimal amount of both energies will be computed under both scenarios. First, regions will act individually in a non-cooperative manner. These results will help us to understand when regions will participate in an International Environmental Agreement, as a second scenario, being required to compare this result with the one obtained when they become signatory regions.

3.4 Non-cooperative Equilibria

Consider that regions act in a non-cooperative scenario. Under this situation, the optimal pair of both energies of a country i, $\{x_i^{NC}, y_i^{NC}\}$, will be the one that solves the following maximization problem:

$$\begin{aligned} \max_{x_i, y_i, \theta_i} \quad & B_i(x_i, y_i) - C(x_i) - K(y_i) - D(E_N) \\ \text{s.t.} \quad & x_i = (\theta_i R)^{\phi}, \\ & y_i \geq 0, \\ & 1 \geq \theta_i \geq 0 \end{aligned}$$

Getting the reduced form of the previous maximization problem by substituting the dirty energy in the production function, one can identify four potential scenarios from the *Kuhn-Tucker* conditions[6]:

Case I Suppose neither dirty nor clean energy is produced, $\theta_i^{NC} = 0$ (equivalent to $x_i^{NC} = 0$) and $y_i^{NC} = 0$. Then, this cannot be an equilibrium since increasing marginally either dirty or clean energy, the country will obtain higher output since $a_i^x \geq d\delta$ and $a^y > 0$, respectively.

Case II Suppose only clean energy is produced, $\theta_i^{NC} = 0$ (equivalent to $x_i^{NC} = 0$) and $y_i^{NC} > 0$. Due to the imperfect substitution between both types of energy, a country could achieve the equilibrium under this situation if the implication of increasing marginally the dirty fossil from 0 will not induce

---

[6] See Proof 6, for more detail.

a positive profit: $a_i^x < \frac{a^y}{1+k} + d\delta$. When previous feature is satisfied for a certain country i, their production level will be: $x_i^{NC} = 0$ and $y_i^{NC} = \frac{a^y}{(1+k)}$.

Case III Suppose only dirty energy is produced, $\theta_i^{NC} > 0$ (equivalent to $x_i^{NC} > 0$) and $y_i^{NC} = 0$. This situation cannot be an equilibrium since countries have incentives to produce and consume clean energy, because it is an imperfect substitute for dirty energy and it is preferred over other types of energies, $a^y > a_i^x$. Both characteristics induce the country to produce at least some amount of clean energy since it is profitable compared to dirty ones.

Case IV Suppose both types of energy are produced, $\theta_i^{NC} > 0$ (equivalent to $x_i^{NC} > 0$) and $y_i^{NC} > 0$. There is an equilibrium depending on the characteristics of the country:

- If the available resources are sufficiently lower, $R \leq \left(\frac{(1+k)\left(a_i^x - d\delta\right) - a^y}{((1+k)(1+c)-1)}\right)^{\frac{1}{\phi}}$; then country i will fully exhaust the resources, producing $x_i^{NC} = R^\phi$ and $y_i^{NC} = \frac{a^y - R^\phi}{(1+k)}$.
- If the available resources are higher enough, $R > \left(\frac{(1+k)\left(a_i^x - d\delta\right) - a^y}{((1+k)(1+c)-1)}\right)^{\frac{1}{\phi}}$; then country i will partially exhaust the resources, producing $x_i^{NC} = \frac{(1+k)\left(a_i^x - d\delta\right) - a^y}{(1+k)(1+c)-1}$ and $y_i^{NC} = \frac{(1+c)a^y - \left(a_i^x - d\delta\right)}{(1+k)(1+c)-1}$.

From the previous explanation, we can summarize that countries will produce the following optimal amount of energy conditional on their own characteristics:

$$x_i^{NC} = \begin{cases} R^\phi, & \text{if } R \leq \left(\frac{(1+k)\left(a_i^x - d\delta\right) - a^y}{((1+k)(1+c)-1)}\right)^{\frac{1}{\phi}} \text{ and } a_i^x \geq \frac{a^y}{1+k} + d\delta \\ \frac{(1+k)\left(a_i^x - d\delta\right) - a^y}{(1+k)(1+c)-1}, & \text{if } R > \left(\frac{(1+k)\left(a_i^x - d\delta\right) - a^y}{((1+k)(1+c)-1)}\right)^{\frac{1}{\phi}} \text{ and } a_i^x \geq \frac{a^y}{1+k} + d\delta \\ 0, & \text{if } a_i^x < \frac{a^y}{1+k} + d\delta \end{cases}$$

And:

$$y_i^{NC} = \begin{cases} \frac{a^y - R^\phi}{(1+k)}, & \text{if } R \leq \left(\frac{(1+k)\left(a_i^x - d\delta\right) - a^y}{((1+k)(1+c)-1)}\right)^{\frac{1}{\phi}} \text{ and } a_i^x \geq \frac{a^y}{1+k} + d\delta \\ \frac{(1+c)a^y - \left(a_i^x - d\delta\right)}{(1+k)(1+c)-1}, & \text{if } R > \left(\frac{(1+k)\left(a_i^x - d\delta\right) - a^y}{((1+k)(1+c)-1)}\right)^{\frac{1}{\phi}} \text{ and } a_i^x \geq \frac{a^y}{1+k} + d\delta \\ \frac{a^y}{(1+k)}, & \text{if } a_i^x < \frac{a^y}{1+k} + d\delta \end{cases}$$

Notice that the optimal amount of energy under non-cooperation will depend on the individual characteristics of the countries, a_i^x. Notice that there are few countries, characterized by $a_i^x < \frac{a^y}{1+k} + d\delta$, that they will focus all their energy in the production of dirty energy $x_i^{NC} = 0$ and $y_i^{NC} = \frac{a^y}{(1+k)}$. The rest will produce both types of energy simultaneously. However, there will be some countries that will partially exhaust their dirty energy resources if $R > \left(\frac{(1+k)\left(a_i^x - d\delta\right) - a^y}{((1+k)(1+c)-1)}\right)^{\frac{1}{\phi}}$, producing and consuming $x_i^{NC} = \frac{(1+k)\left(a_i^x - d\delta\right) - a^y}{(1+k)(1+c)-1}$ and $y_i^{NC} = \frac{(1+c)a^y - \left(a_i^x - d\delta\right)}{(1+k)(1+c)-1}$. Others, with $R \leq \left(\frac{(1+k)\left(a_i^x - d\delta\right) - a^y}{((1+k)(1+c)-1)}\right)^{\frac{1}{\phi}}$, will fully exhaust their dirty energy resources, reaching $x_i^{NC} = R^\phi$ and $y_i^{NC} = \frac{a^y - R^\phi}{(1+k)}$.

Previous levels will determine the aggregate emissions level, E^{NC}, equivalent to:

$$E^{NC} = \frac{R^\phi}{a^y - d\delta}\left[(a^y - \tilde{a}_{NC}^x) + \frac{[(1+k)(1+c)-1]}{2(1+k)}R^\phi\right]N$$

where $\tilde{a}_{NC}^x = \frac{[(1+k)(1+c)-1]R^\phi + a^y}{(1+k)} + d\delta$ is the lowest level of a_i^x that a country fully exhausts their resources under no cooperation.

Once we have determined the non-cooperative equilibria, the next step will focus on the outcomes obtained from the International Environmental Agreement. Additionally, this equilibrium will help us to determine the country decision of participating in the cooperation, being essential in the understanding of the stable coalition concept.

3.5 Cooperative Equilibria

As mentioned previously, the International Environmental Agreement consists of three different stages. First, countries decide whether to join or not the coalition. Once they have joined, and become signatories, they will perceive a positive impact on their welfare function, represented by the contribution to the increase of clean energy production induced by the cooperation through the pro-environmental parameter, β. This coalition game forces signatory regions to choose collectively the maximum percentage of resources to use, $\bar{\theta}^{IEA}$. Since it is restricted between 0 and 1, it is important to consider what occurs under: *no intervention*, *Phase Out*, and *Phase Down*. As a last stage, according to the agreed percentage of resource use, countries will decide the energy production level in a non-cooperative manner.

By using backward induction, the signatory countries will act in a non-cooperative manner at the last third stage, facing the following optimization problem:

$$\begin{aligned} \max_{x_j, y_j, \theta_j} \quad & B_j(x_j, y_j) + \beta \textstyle\sum_{j=1}^{M} y_j - C(x_j) - K(y_j) - D(E_N) \\ \text{s.t.} \quad & x_j = (\theta_j R)^{\phi}, \\ & y_j \geq 0, \\ & \bar{\theta}^{IEA} \geq \theta_j \geq 0 \end{aligned}$$

As done in the previous section, once the previous maximization problem is reduced by substituting the dirty energy constraint by the production function, four potential scenarios appeared from the *Kuhn-Tucker* conditions[7]:

Case I Suppose neither dirty nor clean energy is produced in the signatory country j, $\theta_j^{IEA} = 0$ (equivalent to $x_j^{IEA} = 0$) and $y_j^{IEA} = 0$. Since increasing marginally either dirty or clean energy, the country will obtain higher output since $a_j^x \geq d\delta$ and $a^y + \beta > 0$, respectively; then, this scenario cannot be an equilibrium.

Case II Suppose only clean energy is produced in signatory country j, $\theta_j^{IEA} = 0$ (equivalent to $x_j^{IEA} = 0$) and $y_j^{IEA} > 0$. Again, due to the imperfect substitution between both types of energy, a signatory country could achieve the equilibrium under this situation when the production of dirty energy implies a non-positive profit: $a_j^x < \frac{a^y + \beta}{1+k} + d\delta$. Those signatory countries that satisfy the previous condition will not have any incentive to produce dirty energy due to their reduced impact on the welfare function compared with clean ones. Under this situation, the production level will be: $x_j^{IEA} = 0$ and $y_j^{IEA} = \frac{a^y + \beta}{(1+k)}$.

Case III Suppose only dirty energy is produced in signatory country j, $\theta_j^{IEA} > 0$ (equivalent to $x_j^{IEA} > 0$) and $y_j^{IEA} = 0$. This situation cannot be an equilibrium due to the implication of clean energy in the welfare function, $a^y > a_i^x$, and the imperfect substitution characterization between both types of energies. Both factors will incentive the signatory country to produce at least some amount of clean energy.

Case IV Suppose both types of energy are produced in the signatory country j, $\theta_j^{IEA} > 0$ (equivalent to $x_j^{IEA} > 0$) and $y_j^{IEA} > 0$. There is an equilibrium depending on the characteristics of the signatory country:

---

[7] See Proof 7, for more details.

- If the imposed available resources are sufficiently lower:

$$\bar{\theta}^{IEA}R \leq \left(\frac{(1+k)\left(a_i^x - d\delta\right) - \left(a^y + \beta\right)}{((1+k)(1+c)-1)}\right)^{\frac{1}{\phi}}$$

Then, the signatory country j will fully exhaust the resources:

$$x_j^{IEA} = \left(\bar{\theta}^{IEA}R\right)^{\phi}$$

And:

$$y_j^{IEA} = \frac{\left(a^y + \beta\right) - \left(\bar{\theta}^{IEA}R\right)^{\phi}}{(1+k)}$$

.

- If the imposed available resources are higher enough:

$$\bar{\theta}^{IEA}R > \left(\frac{(1+k)\left(a_i^x - d\delta\right) - \left(a^y + \beta\right)}{((1+k)(1+c)-1)}\right)^{\frac{1}{\phi}}$$

Then, the signatory country j will partially exhaust the resources:

$$x_j^{IEA} = \frac{(1+k)\left(a_i^x - d\delta\right) - \left(a^y + \beta\right)}{(1+k)(1+c)-1}$$

And:

$$y_j^{IEA} = \frac{(1+c)\left(a^y + \beta\right) - \left(a_i^x - d\delta\right)}{(1+k)(1+c)-1}$$

As it will be checked later under the situation of no intervention, the International Environmental Agreement allows countries to produce optimally more clean energy in detriment of dirty ones. From the previous procedure, the signatory countries will produce the following optimal level of each type of energy:

$$x_j^{IEA}\left(\bar{\theta}^{IEA}\right) = \begin{cases} \left(\bar{\theta}^{IEA}R\right)^{\phi}, & \text{if } R < \dfrac{\left(\frac{(1+k)\left(a_j^x - d\delta\right) - \left(a^y+\beta\right)}{(1+k)(1+c)-1}\right)^{\frac{1}{\phi}}}{\bar{\theta}^{IEA}} \text{ and } a_j^x \geq \frac{a^y+\beta}{1+k} + d\delta \\ \frac{(1+k)\left(a_j^x - d\delta\right) - \left(a^y+\beta\right)}{(1+k)(1+c)-1} & \text{if } R \geq \dfrac{\left(\frac{(1+k)\left(a_j^x - d\delta\right) - \left(a^y+\beta\right)}{(1+k)(1+c)-1}\right)^{\frac{1}{\phi}}}{\bar{\theta}^{IEA}} \text{ and } a_j^x \geq \frac{a^y+\beta}{1+k} + d\delta \\ 0, & \text{if } a_j^x < \frac{a^y+\beta}{1+k} + d\delta \end{cases} \tag{3.1}$$

And:

$$.y_j^{IEA}\left(\bar{\theta}^{IEA}\right) = \begin{cases} \frac{\left(a^y+\beta\right) - \left(\bar{\theta}^{IEA}R\right)^{\phi}}{(1+k)}, & \text{if } R < \dfrac{\left(\frac{(1+k)\left(a_j^x - d\delta\right) - \left(a^y+\beta\right)}{(1+k)(1+c)-1}\right)^{\frac{1}{\phi}}}{\bar{\theta}^{IEA}} \text{ and } a_j^x \geq \frac{a^y+\beta}{1+k} + d\delta \\ \frac{(1+c)\left(a^y+\beta\right) - \left(a_j^x - d\delta\right)}{(1+k)(1+c)-1}, & \text{if } R \geq \dfrac{\left(\frac{(1+k)\left(a_j^x - d\delta\right) - \left(a^y+\beta\right)}{(1+k)(1+c)-1}\right)^{\frac{1}{\phi}}}{\bar{\theta}^{IEA}} \text{ and } a_j^x \geq \frac{a^y+\beta}{1+k} + d\delta \\ \frac{a^y+\beta}{(1+k)}, & \text{if } a_j^x < \frac{a^y+\beta}{1+k} + d\delta \end{cases} \tag{3.2}$$

Notice that the optimal amount of energy under cooperation will depend on the agreed maximum percentage of dirty energy resources used, $\bar{\theta}^{IEA}$, chosen in the previous stage. Specifically, this parameter will reduce the amount of legal inputs to exhaust in the signatory country, increasing the chances of fully exhausting their dirty energy reserves at the percentage imposed by the coalition. However, when they do not have incentives to fully exhaust all their dirty energy resources, signatory countries will produce $x_j^{IEA} = \frac{(1+k)\left(a_j^x - d\delta\right) - (a^y+\beta)}{(1+k)(1+c)-1}$ and $y_j^{IEA} = \frac{(1+c)(a^y+\beta) - \left(a_j^x - d\delta\right)}{(1+k)(1+c)-1}$. Previous behavior belongs to those signatory countries such that their preference parameter a_j^x is sufficiently lower. For the rest of other countries, where $a_j^x \geq \frac{a^y+\beta}{1+k} + d\delta$, their energy will be based only on the renewable ones, specifically $y_j^{IEA} = \frac{a^y+\beta}{(1+k)}$. As we will observe later, the presence of the preference parameter, β, will increase the incentives of signatory countries to reduce, or even stop, the extraction of dirty energy, by affecting their optimal level or the threshold to stop

Once identified the equilibrium in the third stage, the signatory countries will decide which is the maximum percentage of dirty energy resources to extract that reduces the aggregate damage level. This decision will affect the early stage, where countries decide whether to join or not the coalition. In order to perform a thorough analysis, different scenarios will be under review, considering the extremal scenarios, *no intervention* ($\bar{\theta}^{IEA} = 1$) and *Phase Out* ($\bar{\theta}^{IEA} = 0$); and the interior potential solutions that encompass the *Phase Down*, where $\bar{\theta}^{IEA} \in (0,1)$.

3.5.1 No intervention

Consider that all signatory countries decide to not intervene in the maximum amount of resource to extract, setting $\bar{\theta}^{IEA} = 1$. This scenario will provide the sterilized effect of the pro-environmental parameter, β, since there is no influence in the maximum percentage level of dirty energy extraction, compared with the non-cooperative scenario. Under this situation, the signatory regions will set the following dirty and clean energy production levels:

$$x_j^{IEA}\left(\bar{\theta}^{IEA} = 1\right) = \begin{cases} R^{\phi}, & \text{if } R < \left(\frac{(1+k)\left(a_j^x - d\delta\right) - (a^y+\beta)}{(1+k)(1+c)-1}\right)^{\frac{1}{\phi}} \text{ and } a_j^x \geq \frac{a^y+\beta}{1+k} + d\delta \\ \frac{(1+k)\left(a_j^x - d\delta\right) - (a^y+\beta)}{(1+k)(1+c)-1} & \text{if } R \geq \left(\frac{(1+k)\left(a_j^x - d\delta\right) - (a^y+\beta)}{(1+k)(1+c)-1}\right)^{\frac{1}{\phi}} \text{ and } a_j^x \geq \frac{a^y+\beta}{1+k} + d\delta \\ 0, & \text{if } a_j^x < \frac{a^y+\beta}{1+k} + d\delta \end{cases}$$

And:

$$y_j^{IEA}\left(\bar{\theta}^{IEA}=1\right)=\begin{cases}\frac{(a^y+\beta)-R^{\phi}}{(1+k)}, & \text{if } R<\left(\frac{(1+k)\left(a_j^x-d\delta\right)-(a^y+\beta)}{(1+k)(1+c)-1}\right)^{\frac{1}{\phi}} \text{ and } a_j^x\geq\frac{a^y+\beta}{1+k}+d\delta\\ \frac{(1+c)(a^y+\beta)-\left(a_j^x-d\delta\right)}{(1+k)(1+c)-1}, & \text{if } R\geq\left(\frac{(1+k)\left(a_j^x-d\delta\right)-(a^y+\beta)}{(1+k)(1+c)-1}\right)^{\frac{1}{\phi}} \text{ and } a_j^x\geq\frac{a^y+\beta}{1+k}+d\delta\\ \frac{a^y+\beta}{(1+k)}, & \text{if } a_j^x<\frac{a^y+\beta}{1+k}+d\delta\end{cases}$$

Notice that, as shown in the following figure, joining the coalition will allow countries to weakly reduce the extraction of dirty energy, but expand the production of clean ones:

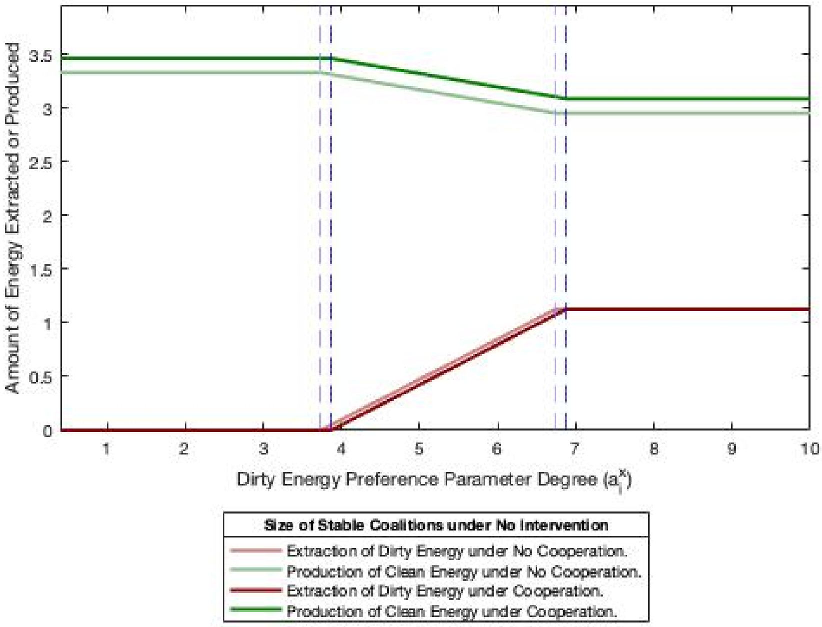


Figure 3.1: Variation of the extraction and production of both types of energy before and after joining the International Environmental Agreement.

Figure 3.1 compares the amount of dirty (*red lines*) and clean energy (*green lines*) extracted, or produced, depending on the dirty energy preference parameter degree, a_i^x, before (*light*) and after (*dark*) joining the International Environmental Agreement. The *two left blue dashed lines* represent the fossil fuel preference parameter threshold where a country decides to move from null to partial exhaustion of dirty energy resources. The *two right blue dashed* lines represent the dirty energy preference parameter threshold where a country decides to move from partial to full exhaustion of dirty energy resources. The *light blue dashed line* represents the thresholds in the non-cooperative scenario, whereas the *dark blue dashed line* represents the threshold in the cooperative scenario.

It is clearly stated that, under no intervention, the presence of the pro-environmental preference helps countries to reduce the level of emissions by the substitution effect generated from dirty energies to their alternative ones. The reduction of the non-renewable energy will occur to a specific range of countries characterized by $a_i^x \in \left(\frac{a^y}{1+k} + d\delta, \frac{[(1+c)(1+k)-1]R^{\psi}+(a^y+\beta)}{1+k} + d\delta\right)$, where the increase in the pro-environmental preference parameter motivates them to reduce the amount of dirty energy. Specifically, few countries, characterized by $a_i^x \in \left(\frac{a^y}{1+k} + d\delta, \frac{a^y+\beta}{1+k} + d\delta\right)$, will stop the extraction of dirty energy as a response of the incentives generated from the International Environmental Agreements. Others, characterized by $a_i^x \in \left(\frac{a^y+\beta}{1+k} + d\delta, \frac{[(1+c)(1+k)-1]R^{\psi}+(a^y+\beta)}{1+k} + d\delta\right)$, will switch some unites of dirty to clean energy as a response of the pro-environmental behavior of the cooperation. There is an extremal case, when $\beta > \left[(1+k)(1+c)-1\right] R^{\phi}$, where countries, that fully exhaust their dirty energy sources, will stop all their extraction to produce only clean energy.

Due to the influence of joining the International Environmental Agreement, countries will become signatories if, once they joined to the cooperation, their welfare level weakly improves compared to the non-cooperative scenario. According to the optimal decision of countries under no intervention in the International Environmental Agreement and the positive effect of the pro-environmental behavior parameter on the welfare function, $\frac{\partial W_j^{IEA}(a_j^x|\bar{\theta}^{IEA},y_j)}{\partial\beta} > 0$, they will contribute to the formation of the grand coalition as the stable coalition under such situation[8]. Since all N countries have incentives to become signatories due to the pro-environmental behavior induced by the International Environmental Agreement. Additionally, it will contribute to the reduction in the aggregate damage level:

$$E^{\bar{\theta}^{IEA}=1} = \frac{R^{\phi}}{a^y - d\delta}\left[\left[a^y - \tilde{a}^x_{\bar{\theta}^{IEA}=1}\right] + \frac{[(1+k)(1+c)-1]}{2(1+k)}R^{\phi}\right]N$$

where $\tilde{a}^x_{\bar{\theta}^{IEA}=1} = \frac{[(1+k)(1+c)-1]R^{\phi}+(a^y+\beta)}{(1+k)} + d\delta$ is the lowest level of a_i^x that a country fully exhausts their resources under a coalition with no intervention.

Notice that a higher pro-environmental parameter reduces the aggregate level of emissions since $\frac{\partial E^{\bar{\theta}^{IEA}=1}}{\partial\beta} = \frac{\partial E^{\bar{\theta}^{IEA}=1}}{\partial\tilde{a}^x_{\bar{\theta}^{IEA}=1}}\frac{\partial\tilde{a}^x_{\bar{\theta}^{IEA}=1}}{\partial\beta} < 0$. In addition, the aggregate level of emissions under no intervention will be reduced when it is compared with the one obtained under no cooperation, $E^{\bar{\theta}^{IEA}=1} < E^{NC}$, at any $\beta > 0$.

The International Environmental Agreement contributes to the more use of clean energy due to its pro-environmental behavior, dropping the incentives of extracting dirty energy. As a consequence, the aggregate emissions level reduces since the incentive to

---

[8] See Proof 8, for a detailing explanation about the decisions of countries to join the International Environmental Agreement.

extract dirty energy will diminish and less countries will fully exhaust their available resources.

3.5.2 Phase Out

Now, consider that the coalition does not accept any extraction of the polluting source, being $\bar{\theta}^{IEA} = 0$. At this stage, the signatory countries will maximize the welfare function at only clean energy levels, being the optimal amount of both energies $x_j^{IEA}(\bar{\theta}^{IEA} = 0) = 0$ and $y_j^{IEA}(\bar{\theta}^{IEA} = 0) = \frac{a^y+\beta}{1+k}$.

Since the coalition has decided on the no availability of dirty energy at the second stage, regions will participate in the International Environmental Agreement if the outcome of such coalition weakly exceeds the one obtained by no cooperating. As proved in Proof 9 and shown in the following figure, the resulting stable coalition will depend positively on the pro-environmental parameter, since it will help the signatory countries to substitute the dirty energy for clean one:

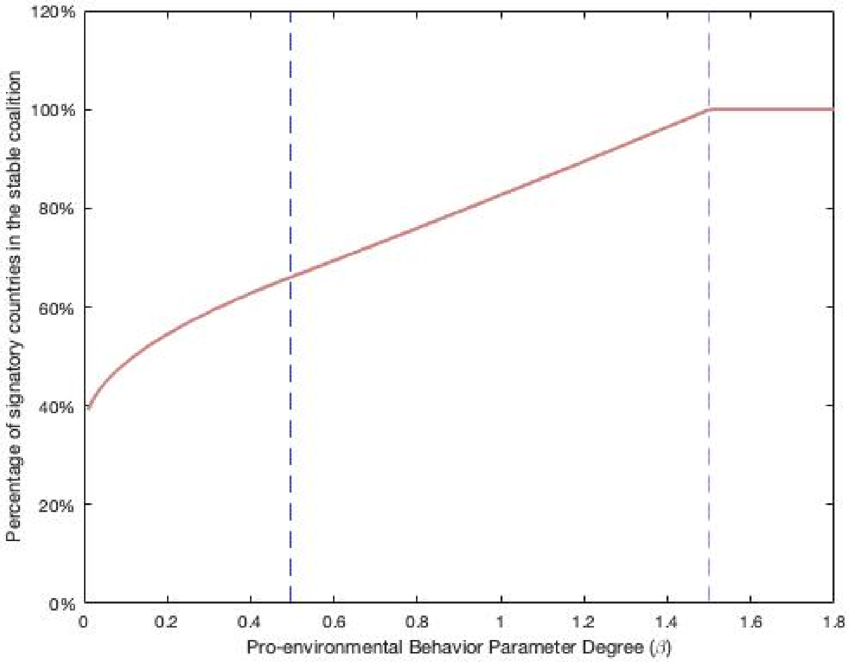


Figure 3.2: Size, in percentage, of signatory countries in the stable coalition under Phase Out at different pro-environmental behavior parameter levels, β.

Figure 3.2 shows the percentage of countries in the stable coalition under *Phase Out* at different pro-environmental behavior parameter levels, β. The left dark blue dashed line indicates when countries that fully exhaust resources prior to the International Environmental Agreement join the coalition. The right light blue dashed line indicates the highest pro-environmental preference parameter such that all countries become signatories and form the grand coalition.

As a first observation, the countries with sufficient lower characterization a_i^x, less than $\frac{a^y}{1+k}+d\delta$, will join the coalition, since the presence of the pro-environmental parameter will strongly improve their welfare function. However, the decision of the rest of the countries will depend on the level of β, since it will compensate for the loss generated by not extracting more dirty energy. Following this reasoning, the effect of increasing the pro-environmental preference parameter on the number of countries in the stable coalition will be positive, since at a higher parameter, signatory regions more dependent on dirty energy, at higher characterization a_i^x, will be able to offset the loss generated by stopping extraction in their dirty energy sources. Notice that those countries that before fully exhaust their resources in the International Environmental Agreement, at $\beta \approx 0.5$ (left dark blue dashed line in the figure), will join the coalition even if there is a *Phase Out*. There is a certain threshold, $\beta > 1.5$ (right light blue dashed line in the figure), such that, once exceeded, will form the grand coalition.

Given the stable coalition identified before, the aggregate level of emissions will be:

$$E^{\bar{\theta}^{IEA}=0} = \begin{cases} 0, & \text{if } \beta \geq \beta^{''} \\ \frac{R^{\phi}\left[a^y - a^{x''}(\beta)\right]}{a^y - d\delta}, & \text{if } \beta \in [\beta^{'}, \beta^{''}) \\ \frac{R^{\phi}\left[a^y - \tilde{a}^x_{NC}\right]}{a^y - d\delta} + \left[\frac{(1+k)(a_i^x)^2 - 2\left(a_i^x\right)[(1+k)d\delta + (a^y+\beta)]}{2[(1+k)(1+c)-1](a^y - d\delta)}\right]_{a_i^x = a^{x'}(\beta)}^{a_i^x = \tilde{a}^x_{NC}}, & \text{if } \beta \in (0, \beta^{'}) \end{cases}$$

where $\beta^{'}$ and $\beta^{''}$ corresponds to the thresholds[9] [10] that all ex-ante partial and full exhauster, respectively, improves their welfare under the *Phase Out* and accepts to join the International Environmental Agreement. Additionally, $a^{x'}(\beta)$ and $a^{x''}(\beta)$ represents the fossil fuel preference parameter of the last partial and last full exhauster country that will join the International Environmental Agreement, respecting $W_j^{IEA}\left(a^{x'}, \beta | \theta_j = 0, y_j > 0\right) = W_i^{NC}\left(a^{x'} | \theta_j \in (0,1), y_j > 0\right)$and $W_j^{IEA}\left(a^{x''}, \beta | \theta_j = 0, y_j > 0\right) =$

---

[9] Where $\beta^{'}$ is equivalent to:

$$\sqrt{(1+k)\frac{[(1+k)(\tilde{a}^x_{NC} - d\delta)^2 + (1+c)(a^y)^2 - 2a^y(\tilde{a}^x_{NC} - d\delta)]}{[(1+k)(1+c)-1]}} - a^y$$

Being $\tilde{a}^x_{NC} = \frac{[(1+k)(1+c)-1]R^{\phi} + a^y}{(1+k)} + d\delta$.

[10] Where $\beta^{''}$ is equivalent to:

$$\sqrt{\frac{(1+k)\left[(1+k)(a^y - d\delta)(\tilde{a}^x_{NC} - d\delta) + (1+c)(a^y)^2 - (1+k)(\tilde{a}^x_{NC} - d\delta)\left[\tilde{a}^x_{NC} - a^y\right] - 2a^y(a^y - d\delta)\right]}{[(1+k)(1+c)-1]}} - a^y$$

Being $\tilde{a}^x_{NC} = \frac{[(1+k)(1+c)-1]R^{\phi} + a^y}{(1+k)} + d\delta$.

$W_i^{NC}\left(a^{x''}|\theta_j = 1, y_j > 0\right)$, respectively.

Under *Phase Out* the number of signatory countries in the stable coalition will depend on the pro-environmental behavior parameter. In other words, a grand coalition is possible to be formed when the previous parameter is higher enough. Although *Phase Out* retains countries to join the International Environmental Agreement, it reduces the emissions compared to the non-cooperative scenario when the pro-environmental behavior parameter increases.

3.5.3 Phase Down

Once both extremal cases are analysed, the study will focus on any potential result in $\bar{\theta}^{IEA} \in (0,1)$. By using backward induction, the signatory regions will set the dirty and clean energy production in the third stage at the levels determined in the beginning of the current section, Equation 3.1 and 3.2.

Before acting as non-cooperative entities under the influence of the coalition, the signatory regions need to decide the maximum percentage of available dirty resources to extract, $\tilde{\theta}^{IEA}$. In a cooperative manner, all the signatory countries choose the optimal rate such that the worldwide emissions level will be reduced subject to the participation constraint of all their members:

$$\begin{aligned} \min_{\theta} \quad & D(E^{\theta^{IEA}\in(0,1)}) = d\textstyle\sum_{i=1}^{N} e_i = d\delta \textstyle\sum_{i=1}^{N} x_i\left(\theta\right) \\ \\ \text{s.t.} \quad & \theta \in (0,1) \\ & W_j^{IEA}\left(a_j^x|\theta \in (0,1), y_j > 0\right) \geq W_j^{NC}\left(a_j^x|\theta_j \in [0,1], y_j > 0\right), \quad \forall j \in M \end{aligned}$$

In other words, the signatory countries search for the maximum percentage of available dirty resources to extract that minimizes the aggregate damage level, respecting the participation constraint of the signatory countries to accept the International Environmental Agreement. Notice that all these constraints can be transformed into a new function, Δ_W, that computes the difference between the welfare levels after and before joining the International Environmental Agreement:

$$\Delta_W(a_j^x, \vartheta, \beta) = W_j^{IEA}\left(a_j^x|\vartheta \in (0,1), y_j > 0\right) - W_j^{NC}\left(a_j^x|\theta_j \in [0,1], y_j > 0\right)$$

By evaluating no difference between welfare levels, $\Delta_W = 0$, the Implicit Function Theorem will be used to observe the potential maximum percentage of available dirty resources to extract under increasing levels of the dirty energy preference parameter:

$$\frac{\partial \vartheta}{\partial a_j^x} = -\frac{\frac{\partial \Delta_W}{\partial a_j^x}}{\frac{\partial \Delta_W}{\partial \vartheta}}$$

The previous expression will explain the effect of joining in the coalition the country with a higher dirty energy preference parameter, a^x, on the maximum percentage of available dirty resources to extract. In other words, it shows how much the maximum percentage of available dirty resources to extract should increase in order to accept a country with a marginally higher level of dirty energy preference parameter. Starting with the effect of θ on Δ_W, $\frac{\partial \Delta_W}{\partial \theta}$:

$$\frac{\partial \Delta_W}{\partial \theta} = \frac{\partial W_j^{IEA}\left(a_j^x | \theta \in (0,1), y_j > 0\right)}{\partial \theta} =$$

$$= \phi \frac{(\theta R)^{\phi}}{\theta}\left[(a_j^x - d\delta) - (1+c)(\theta R)^{\phi} - \frac{(a^y+\beta) - (\theta R)^{\phi}}{(1+k)}\right] - \frac{\partial \sum_{-i \in N/\{j\}} x_{-i}}{\partial \theta} > 0$$

This first partial derivative[11] comes from two different facts. First, the positive impact on the individual welfare (expressed in the first term), since the reduction of the maximum percentage of available dirty resources to extract will restrict the amount of dirty energy to use and, at the same time, will increase the renewable one. Notice that the reduction of θ will affect those countries that fully exhaust all their available resources since $(\theta R)^{\phi} = \frac{(1+k)\left(\tilde{a}_{PD}^x - d\delta\right) - (a^y+\beta)}{(1+k)(1+c)-1}$ and $a_j^x > \tilde{a}_{PD}^x = \frac{[(1+k)(1+c)-1](\theta R)^{\phi} + (a^y+\beta)}{(1+k)} + d\delta$. Second, an increase in the maximum percentage of available dirty resources to extract will imply a reduction in the level of emissions, $\frac{\partial \sum_{-i \in N/\{j\}} x_{-i}}{\partial \vartheta} < 0$, since more countries will join to the International Environmental Agreement, perceiving the impact of the pro-environmental behavior parameter and reducing the dirty energy extraction.

Now, let's consider the effect of the dirty energy preference parameter on the country welfare level:

$$\frac{\partial \Delta_W}{\partial a_j^x} = \frac{\partial W_j^{IEA}\left(a_j^x | \theta \in (0,1), y_j > 0\right)}{\partial a_j^x} - \frac{\partial W_j^{NC}\left(a_j^x | \theta_j \in [0,1], y_j > 0\right)}{\partial a_j^x}$$

Depending on the scenario that signatory region j were under non-cooperation, previous differentiation could take two different forms. If $R^{\phi} > \frac{(1+k)\left(a_i^x - d\delta\right) - a^y}{((1+k)(1+c)-1)}$:

$$\frac{\partial \Delta_W}{\partial a_j^x} = (\theta R)^{\phi} - \frac{(1+k)(a_j^x - d\delta) - a^y}{(1+k)(1+c)-1} < 0$$

---

[11] Since $W_j^{IEA}\left(a_j^x | \vartheta \in (0,1), y_j > 0\right) = (a_j^x - d\delta)(\vartheta R)^{\phi} - (1+c)\frac{(\vartheta R)^{2\phi}}{2} + \frac{\left((a^y+\beta) - (\vartheta R)^{\phi}\right)^2}{2(1+k)} - d\delta \sum_{-i \in N/\{j\}} x_{-i}$.

Since the highest value of $(\theta R)^{\phi}$ is equivalent to $\frac{(1+k)(\tilde{a}^x_{PD}-d\delta)-(a^y+\beta)}{(1+k)(1+c)-1}$ where $a^x_j \geq \tilde{a}^x_{PD}$. Same result would occur if $R^{\phi} \leq \frac{(1+k)\left(a^x_i-d\delta\right)-a^y}{((1+k)(1+c)-1)}$:

$$\frac{\partial \Delta_W}{\partial a^x_j} = \left[\theta^{\phi} - 1\right] R^{\phi} < 0$$

Because joining the International Environmental Agreement under a *Phase Down*, countries with a higher characterization of a^x will perceive a less increase in their welfare function.

According to the previous explanation, $\frac{\partial \vartheta}{\partial a^x_j} > 0$, specifying that the coalition needs to increase the maximum percentage of available dirty resources to extract in order to attract countries with a higher characterization of a^x to join the coalition. Additionally, $\frac{\delta \Delta_W}{\delta a^x_j} < 0$ indicates that at a certain θ that convinces a country described by $\hat{a}^x$ to join the coalition, signatory countries with a lower characterization of dirty energy preference $\hat{a}^x \geq a^x_j$ will show even more incentives to join the International Environmental Agreement, since it will result in a greater welfare level, compared to the non-cooperative scenario. From this perspective, binding the participation constraint of the country with the highest a^x_j will respect the participation constraint of the rest of the countries with a lower characterization of a^x.

By considering the previous explanation, the maximization problem becomes as follows:

$$\begin{aligned}
\min_{\theta} \quad & D\left(E^{\theta^{IEA} \in (0,1)}\right) = d\delta N \left[\int_{\mathbf{a^x}}^{a^y} x^{NC}_i \frac{1}{a^y-d\delta} da^x_i + \int_{\tilde{a}^x}^{\mathbf{a^x}} x^{IEA}_i(\theta) \frac{1}{a^y-d\delta} da^x_i + \int_{a^y}^{\tilde{a}^x} x^{IEA}_i(\theta) \frac{1}{a^y-d\delta} da^x_i\right] \\
\text{s.t.} \quad & \theta \in (0,1) \\
& W^{IEA}_j\left(\mathbf{a^x} | \theta \in (0,1), y_j > 0\right) = W^{NC}_j\left(\mathbf{a^x} | \theta_j \in [0,1], y_j > 0\right), \quad \text{where } \mathbf{a^x} \in \operatorname{argmax} a^x_j, \forall j \in M
\end{aligned}$$

where $\mathbf{a^x}$ is the signatory country with the highest dirty energy characterization level and, additionally, it binds the participation constraint. In other words, all signatory countries with $\mathbf{a^x} \geq a^x_j$ will accept the conditions of the International Environmental Agreement. By taking the First Order Condition of the previous minimisation problem, we identify:

$$\frac{x^{NC}_i(\mathbf{a^x}) - (\theta R)^{\phi}}{a^y - d\delta} \frac{\partial \mathbf{a^x}}{\partial \theta} = \frac{\phi (\theta R)^{\phi - 1} R}{a^y - d\delta} (\mathbf{a^x} - \tilde{a}^x) \tag{3.3}$$

From the previous condition, the cooperation will choose the optimal rate of maximum available dirty resources to extract such that equals the marginal increase of emissions with the marginal reduction induced by the International Environmental Agreement. On the left side, it is the presence of the marginal increase of emissions $\left(\frac{x^{NC}_i(\mathbf{a^x})-(\theta R)^{\phi}}{a^y-d\delta}\right)$ for those countries that do not prefer to participate in the cooperation because of the variation of the maximum percentage of available dirty resources to extract $\frac{\partial \mathbf{a^x}}{\partial \theta}$. On the right side, there is the marginal reduction of emissions for those countries that participate in the cooperation, $\frac{\mathbf{a^x} - \tilde{a}^x}{a^y - d\delta}$, and fully exhaust their dirty energy resources at the agreed rate $\big(\phi(\theta R)^{\phi-1}R\big)$.

Additionally, Equation 3.3 shows us the rule to determine the optimal maximum percentage of available dirty resources to extract. As an illustration, suppose that the previous equation is satisfied with a strictly greater inequality ($>$), at a certain θ. This rate will induce a higher increase in the emissions from the country that rejects the International Environmental Agreement than the reduction generated from those than remain. Since the cooperation keeps reducing the aggregate damage level, it will motivate them to increase the maximum percentage of available resources to extract. Notice that this expression reveals the tension to move from *Phase Out* to *Phase Down*, only if the pro-environmental behavior parameter is not high enough. Now, consider that previous Equation 3.3 contains a strict lower inequality ($<$), resulting in a higher reduction of emissions generated from countries that remains in the International Environmental Agreement, against the increase from those that reject to be part from such type of coalition. Under this situation, the cooperation will have incentives to still reduce the maximum percentage of available resources to extract, minimizing even more the current aggregate damage level. Again, this situation reveals the reason why the *no intervention* is not placed against *Phase Down*.

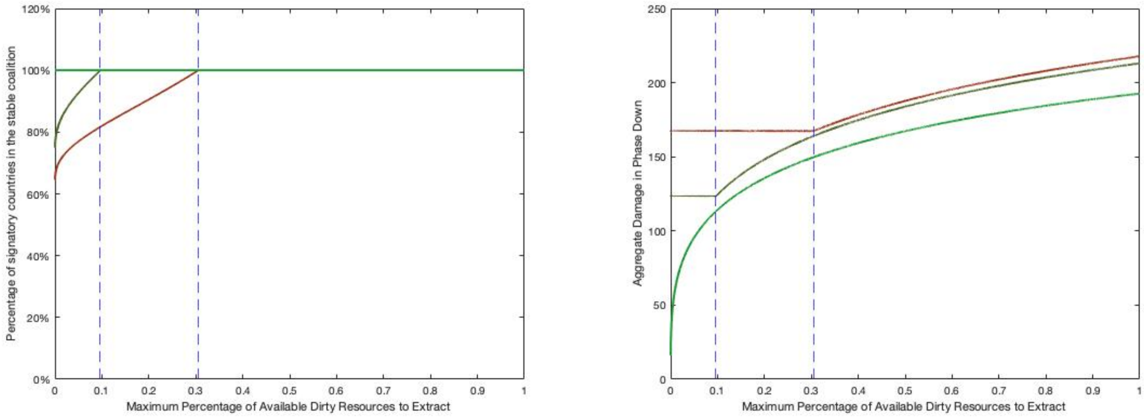
The previous explanation tells the potential mechanism that will allow us to identify the optimal maximum percentage of dirty energy resources to extract:

Proposition 4 *Under Phase Down, the optimal maximal percentage of available resources to extract will be the minimum one that forms the grand coalition.*

Proof: *See in Appendix, Proof 10.*

At the optimal percentage of available dirty energy resources to extract, θ^*, no country will act as a non-cooperative entity due to the lack of incentives, being stable the coalition. Notice that any change will induce an increase in emissions since either a country will leave the International Environmental Agreement or it is allowed to extract more dirty energy.

Now the question is how the pro-environmental preference parameter affects the optimal *Phase Down* decision. Since the decision of entering or not to the coalition affects those countries that fully exhaust their resources, the pro-environmental preference parameter will not influence the decision of either joining or not the coalition. However, the previous parameter will affect directly the profits structure. In other words, as it interacts with the potential benefits obtained by using clean energy; it will make difficult, through the maximal percentage of available resources to extract, the likelihood to become a fully exhauster country. Therefore, higher pro-environmental preference parameters will induce greater efforts to reduce the maximum amount of available dirty energy resources to extract.



(a) Stable Coalition. *(b) Aggregate Damage.*

Figure 3.3: Size, in percentage, of signatory countries in the stable coalition and aggregate damage under Phase Down at different pro-environmental behaviour parameter levels, β.

Figure 3.3 show us the size, in percentage, of the stable coalition and the aggregate damage under *Phase Out* at different pro-environmental behavior parameter levels, respectively. The three lines correspond to a low ($\beta = 0.4 < \beta'$, red one), medium ($\beta' < \beta = 0.7 < \beta''$, brown one), and high ($\beta'' < \beta = 2$, green one) pro-environmental preference parameter, where the dashed line corresponds to the resulting optimal rate chosen by the International Environmental Agreement coalition.

As previously explained, apart from decreasing the aggregate emissions level and, consequently, their damage, the pro-environmental preference parameter helps to reduce the maximum percentage of available dirty resources to extract. Additionally, there is a certain threshold (β''), such that, once exceeded, the cooperation prefer to set the *Phase Out* as an optimal strategy to reduce the emissions level.

The optimal decision of the coalition in the *Phase Down* will depend on the minimization of the aggregate damage subject to the participation constraint of all signatory countries. As a result, the *Phase Down* reduces the aggregate damage compared to the other extremal cases, if the pro-environmental behavior parameter is not high enough. Otherwise, the resulting stable coalition, that captures the core, will establish the *Phase Out*. The *Phase Down* optimization problem could be reduced to identify the maximum percentage of available dirty resources to extract that convinces the grand coalition to be in the International Environmental Agreement. According to the results obtained, it seems that the potential outcome obtained from the *Glasgow Climate Pact* is driven by a lower pro-environmental behavior atmosphere generated in such agreement.

3.6 Conclusion

The *Glasgow Climate Pact* has brought under discussion the management of the unabated coal power. Precisely, the main concern of the bargaining process was to determine the final use of such resources. The asymmetry of countries around the sensibility of fossil fuel in their welfare function complicated the potential coordination of phasing out the unabated coal power. Not only the national characteristics of countries break with such an agreement, but the lack of complete financial aid reduced the incentives of joining *Glasgow Climate Pact*. According to this conflict, the main motivation of this paper was to analyze the main factors that allow the final decision reached in the agreement.

Following the existing literature, our paper has followed the basic linear-quadratic model of an emissions game. Additionally, the current paper considers the approach used in Rubio (2017), where the Kuhn-Tucker conditions become essential to prevent any negative pay-off. In the sense of understanding the size of the coalition, D'Aspremont et al. (1983) is essential to understand the stability conditions to join the agreement. Moreover, the pro-environmental parameter becomes relevant to reproduce the private benefits obtained by the *Glasgow Climate Pact*. The interpretation of this concept was obtained from the *warm-glow* concept idealized by Andreoni (1989, 1990) or the description of the pro-environmental behavior made by Taufik and Venhoeven (2018).

As mentioned previously, the model used to address the question follows the linear-quadratic model of an emissions game. Furthermore, countries are only distinguished by their fossil fuel preference parameter and they decide whether to participate or not in cooperation. If they reject the proposal, they will decide as an individual entity. However, if they decide to join the cooperation, then a bargaining process will occur after perceiving a pro-environmental behavior incentive. First, all signatory countries will decide the maximum percentage of available energy to extract, to reduce the aggregate damage level. Once this rate is decided among all the signatory countries, then they will select individually the amount of both types of energy to spend.

Three different cases are analyzed to understand the International Environmental

Agreement explained before: *no intervention*, *Phase Out*, and *Phase Down.* Starting with the first one, it will reproduce the sterilized effect of joining the coalition without affecting the individual decision extraction. Under this scenario, all countries have incentives to join the coalition if no imposition in the maximum percentage of available resources to extract is established. Due to the effect of the pro-environmental parameter, there is a reduction in the emissions level. Now, consider the hypothetical case that *Phase Out* is established. The potential coalitions will depend on the level of pro-environmental behavior, being possible to form a grand coalition if such a parameter is sufficiently higher. Last, the *Phase Down* will be the process that detects the optimal maximum percentage of available resources such that reduces the aggregate emissions level. The optimization problem will be detected when the marginal change of emissions for the signatory countries equals the marginal variation of emissions for those countries that are acting individually. Apart to analyse that *no intervention* is not an optimal strategy because of the effect of the pro-environmental behavior from joining the International Environmental Agreement, the *Phase Out* is only possible if the pro-environmental behavior is high enough.

According to the previous model, the outcome obtained in the *Glasgow Climate Pact* is the resulting optimal strategy generated by a lower pro-environmental behavior promoted in the 26^{th} *Conference of Parties.* The debate from the countries with a higher pollutants dependency, such as China or India, induces the weak effort made to promote the *Phase Out* among signatory countries. Although the theoretical model provided show us a simple explanation to understand the *Glasgow Climate Pact*, there are many potential extensions of this model. First, the level of resources is assumed to be constant among countries, a fact that is not supposed to be empirically proved due to the difference in sizes among countries. Explaining this scenario will help us to comprehend the effect of *Phase Down* and/or *Out* under other factors. Second, it is assumed a fixed decision for all countries, not allowing the discrimination of strategies. Following this model, it is possible to find a potential solution that could increase further the aggregate welfare, holding the same aggregate damage level. Although both extensions could help us to improve our understanding of the *Glasgow Climate Pact*, it is highlighted in this paper that countries failed to provide more financial programs to support those countries that reduce their welfare because of the non-use of dirty energy resources. According to this statement, the *Phase Down* in the *Glasgow Climate Pact* is the reasonable strategy to reduce the aggregate damage level in the current climate warming situation that we face nowadays.

Conclusion

This thesis comprises of three chapter, one chapter on political economics and two chapters on political and environmental economics. Chapter 1 analyses the constitution of the trust in national parliaments through a game where the population, parties (and politicians), and lobby interact to decide the electoral and the final implemented policy. It describes trust in national parliament as the likelihood that the final implemented policy reaches, at least, the promised level of utility obtained in the electoral competition. From this definition, it studies how trust in national parliament is constituted and which factors can explain its variation. Chapter 2 studies if those demand-side policies that influence the preference for energy are effective in their aim to reduce the potential environmental damage generated from fossil fuels. By using an infinite-horizon non-cooperative differential game, with consumers and multi-supplier firms trading fossil fuel and clean energy, we assess the potential consequences that this type of policy could generate in the environment. Chapter 3 examines the bargaining process behind the *Glasgow Climate Pact* about the state of unabated coal power signed in the last 26^{th} Conference of Parties. Through a basic linear-quadratic model of emissions game, it investigates the main factors and conditions that drive the coalition to *Phase Down* or *Out* their unabated coal power.

Chapter 1 starts with the description of a game that explains how policies are determined in a legislature. There is a timing of eight stages, which include the electoral competition; the changes in the politician state-probability belief; and the lobby contract, which includes the suggested policy and the transfer done to the politician. This sequence explains part of the different pressures that provoke changes from the electoral to the final implemented policies. One, *internal pressure* is related to the difference in the state-probability between population and politician. Another, *external pressure* refers to the potential suggestions that an external agent, lobby, is able to offer to the politician. As a result, the expectations about the politician state-probability or lobby characteristics will affect the constitution of individual trust in national parliament. Specifically, the political difference between most preferred policies among states increases the asymmetry in individual trust in the population by *internal pressures*; whereas the number of individuals, the policy discount parameter, and the politician ratio rise individual trust through *external pressures*.

Chapter 2 contributes to the literature by extending the understanding of the potential implications of how policies addressed to influence the preference for energy impacts the environment. The answer to this research question is identified through a two-periods time model, where two types of energy, brown (polluting) and green (clean) sources, are supplied in the market. In this game, consumers, represented by a parameter that captures the prevalence of clean energy over fossil fuels, and multi-supplier firms, that seek to maximize their profit function, interact to extract or produce both types of energy. By analyzing exogenous shocks in the preference for the use of renewable energy, we identified two different results. First, if there is no incentives to fully exhaust the fossil fuel reserves, more green preferences induce a reduction in the use of fossil fuel sources if and only if the change does not induce the incorporation of renewable energies in the market. However, if there is incentives to fully exhaust the fossil fuel reserves, more green preferences induce an anticipation in the extraction decisions. Additionally, although imperfect competitive markets reduce the potential amount of emissions, perfectly competitive ones seem to reduce the potential augment of emissions when clean energy is newly introduced in the market.

Chapter 3 considers the negotiation process behind the Glasgow Climate Pact to explain the main reason that has led to the determination of the *Phase Down* on the unabated coal power of countries member. The model used is a linear-quadratic model of emissions game, where countries, that are distinguished by their fossil fuel preference parameter, decide whether to join an International Environmental Agreement or not. The coalition promotes the pro-environmental behavior of signatory countries but restricts the use of fossil fuel resources to minimize the aggregate damage. Our findings suggest that *Phase Down* is an optimal cooperative decision if the pro-environmental behavior is lower enough. Otherwise, *Phase Out* is an optimal decision made by the grand coalition, vanishing the potential aggregate damage generated from fossil fuel energy, if the pro-environmental behavior is sufficiently higher.

Now, we can specify the policy implications and potential further research from the three previous chapters. In Chapter 1, our findings suggest that trust in national parliament is sensible to many different factors, captured with the name of *internal* and *external pressures*. Due to the economic and financial volatility of world markets, *internal pressures* are difficult to prevent, but common among countries because of the current globalization. However, the presence of lobbies capable of suggesting, and finally implementing, their political proposal are inevitable. In the end, they can reduce the electoral turnout as a result of the lower credibility of the electoral agenda. Since lobbies are externally motivated to influence the political scenario to maximize their own profits, the potential policy implication from this chapter is to show the need for a lobby that protects the politicians under external suggestions that deviate the final political decision from the internal policy. There is a potential gap in including some empirical research behind the explanation provided in this chapter. Since state-probability is common among countries, it is possible to look for the effect of external pressure on the trust in national parliaments.

The important implication of Chapter 2 is that increasing awareness of the use of clean energy, jointly with the promotion of the affordable renewable energy, could not achieve the goal of reducing the consequences of climate change. Depending on the characteristics of the market, there are two different results obtained. For those economies that fully exhaust their resources, more green preferences could anticipate the extraction decisions. However, it seems that this result is restricted to the limitations of our model, when we do not allow to postpone the extraction decision for future periods. But, for those economies that do not fully exhaust their resources, we have observed that the presence of an alternative in the market could increase the competition between the different types of energies, motivating the anticipation of fossil fuel extraction. This result could generate a conflict with the intention of the *United Nations General Assembly* of reducing the accumulation of CO_2 in the atmosphere. More specific, both 13.3 and 7 SDG could induce an increase in the Net Present Value of Emissions. As a potential motivation from this chapter is that to prevent any consequence of the Green Paradox, policy-makers need to reduce the power of firms in the energy market or the promotion of affordable energy to prevent any indirect effect from inducing the use of clean energy through preferences. As mentioned, there is still research to do around these results, as introducing the concept of asymmetry in imperfectly competitive structures or the presence of many fossil fuel or clean energy alternatives in the energy market.

Chapter 3, for the attention of policymakers, we may argue that the final outcome obtained from the *Glasgow Climate Pact* related to the unabated coal power was driven by the weak pro-environmental behavior promoted in the 26^{th} Conference of Parties. Our argument suggests that the decision of "*accelerating efforts towards the phase-down of unabated coal power*" is influenced by the reduced efforts to provide financial aid to developing countries. One of the potential future research directions is the stability and effects of promoting discrimination in the *Phase Down* and *Out* strategies exerted on the signatory countries. This policy could generate a potential reduction of aggregate damage, allowing an increase in the individual and aggregate welfare levels.

Appendix: Chapter 1

Proof

Proof 1 *Define $F(p)$ as the distribution function of most preferred polices such that, when it is evaluated at p, $F(p)$ indicates the fraction of population whose most preferred polices are less than p. $F(p)$ is an increasing around function since $p' > p$ then $F(p') > F(p)$. Furthermore, $F(p)$ is equivalent to Lebesgue measure, since, for a given two different policies, the set of indifference has F-measure zero.*

Given individual utilities $u_i(p)$ is continuous in their argument and single-peak, for a pair of polices (p_0, p_0), such that $F(p_0) = \frac{1}{2}$, it is a strict Condorcet winner. This result is coming from the fact that for any $p < p_0$, more than one-half of the population prefer p_0 to p. Suppose that it is not true, an one-half the population either prefer p to p_0 or are indifferent between both. Due to the fact that between two different polices, the set of indifferent individuals has F-measure zero, the number of individuals that would prefer either one or the other will be the same. Then, $F(\Omega(p_0, p)) = \frac{1}{2}$, where $\Omega(p_0, p)$ contains the individuals whose most preferred polices are arbitrarily close to p_0.

Consider $\mathcal{I}^0$ as a set of individuals characterized by $\bar{\gamma}_i = p_0$. At any policy $p \neq p_0$, $u_{\mathcal{I}}(p_0) - u_{\mathcal{I}}(p) > 0$. Let $\mathcal{I}_{\mathcal{B}}$ be a set of individuals where their most preferred policy is allocated in an open ball, $\mathcal{B}$, about $\mathcal{I}^0$ such that:

$$i \in \mathcal{I}_{\mathcal{B}} \Rightarrow u_i(p_0) - u_i(p) > 0$$

Select a sequence $\{\mathcal{I}_i\}$ of types in $\Omega(p_0, p)$ whose most preferred policy converge to p_0. In the limit of this sequence, $\mathcal{I}_0$ has most preferred policy p_0. But, it shows that $\mathcal{B}$ about $\mathcal{I}_0$ is constructed in points of $\Omega(p_0, p)$. Then, $F(\Omega(p_0, p)) = \frac{1}{2}$ is false, and majority of population prefer p_0 than p, and p_0 is a strict Condorcet winner (in addition, it is possible to use a similar argument shows when $p > p0$).

From previous definition, we will prove that p_0 is a political equilibrium. Consider next three cases:

1. *There is a policy p_0 such that $F(p_0) = \frac{1}{2}$. Then, pair (p_0, p_0) is a political equilibrium, since no party can obtain profits by deviating their policy. Suppose that a certain party deviate to $p < p_0$. Half of population have their most preferred policies over p_0 and prefer p_0 to p. Then, party gets probability at most one-half for winning, being the deviation non-profitable. Now, suppose that party deviate to $p_0 < p$. Half of*

population have their most preferred policies below p_0 and prefer p_0 to p. Party gets probability at most one-half for winning, being the deviation non-profitable. Then, at (p_0, p_0), there is not profitable deviation. Applying same argument for the other party, show us that (p_0, p_0) is a political equilibrium.

2. *There is no policy p such that $F(p) = \frac{1}{2}$. Suppose $\{p|F(p) \geq \frac{1}{2}\}$ is not empty. If $p_0 = inf\{p|F(p) \geq \frac{1}{2}\}$, then it is a political equilibrium. For example, if a party decides to deviate to $p_0 < p$, half the population prefer p_0 to p, then other party could decide to deviate for any small $\epsilon > 0$, such that $F(p+\epsilon) > \frac{1}{2}$. Due to this strategy, this deviation is unprofitable. But, if a party decides to deviate to $p < p_0$, then party will get $F(p) < \frac{1}{2}$. Due to this fact, this deviation is unprofitable.*
3. *There is no policy such that $\{p|F(p) \geq \frac{1}{2}\} = \emptyset$. Then, $(\bar{p}, \bar{p})$ is an equilibrium, since $F(\bar{p}) \leq \frac{1}{2}$ and more than half the population prefer $\bar{p}$ to any other policy.*

Given that $F(p)$ is continuous and strictly increasing on P, the median ideal policy p_0 results in a strict Condorcet winner, generating a unique political equilibrium in (p_0, p_0).

Proof 2 *Optimal contract solves next maximization problem by binding first constraint:*

$$\begin{aligned} \max_{\{p,t\}} \quad & Y - c\,(\psi - p)^2 - t \\ s.t. \quad & \beta[w + t(p_L)] + (1-\beta)\,E\left[W_P(p)\right] \geq \beta[w] + (1-\beta)\,E\left[W_P(p_P)\right], \\ & Y - c\,(\psi - p)^2 - t(p_L) \geq Y - c\,(\psi - p_P)^2, \\ & p \in [0,1] \end{aligned} \tag{4}$$

By binding first constraint, we obtain that t is equal to:

$$t(p_L) = \frac{(1-\beta)}{\beta}\Big(E\left[W_P(p)\right] - E\left[W_P(p_P)\right]\Big)$$

By substituting previous equation in the objective function and constraint, we obtain:

$$\begin{aligned} \max_{\{p\}} \quad & Y - c\,(\psi - p)^2 - \tfrac{(1-\beta)}{\beta}\Big(E\left[W_P(p)\right] - E\left[W_P(p_P)\right]\Big) \\ s.t. \quad & Y - c\,(\psi - p)^2 - \tfrac{(1-\beta)}{\beta}\Big(E\left[W_P(p)\right] - E\left[W_P(p_P)\right]\Big) \geq Y - c\,(\psi - p_P)^2, \\ & p \in [0,1] \end{aligned} \tag{5}$$

By applying Khun-Tucker, we identify the $\mathcal{L}$agrange function from previous maximization *problem, where there is not binding constraint.*

$$\mathcal{L}(p) = Y - c\left(\psi - p\right)^2 - \frac{(1-\beta)}{\beta}\Big(E\left[W_P(p)\right] - E\left[W_P(p_P)\right]\Big)$$

It could be rewrite as:

$$\mathcal{L}(p) = Y - c\left(\psi - p\right)^2 - \frac{(1-\beta)}{\beta} d \left[\sum_{i=1}^{n}\left\{(p-p_P)^2 + 2(p-p_P)(p_P - \bar{\gamma}_{P,i})\right\}\right]$$

Where $\bar{\gamma}_{P,i} = \pi_P \gamma_N^A + (1-\pi_P)\gamma_N^B$*.Taking F.O.C. from previous function with respect to p:*

$$\frac{\delta \mathcal{L}(p)}{\delta p}: \quad 2c\left(\psi - p\right) + \frac{(1-\beta)}{\beta} 2d \sum_{i=1}^{n}(\bar{\gamma}_{P,i} - p) = 0$$

Or, rearranging and considering p as a constant and $\sum_{i=1}^{n}(\bar{\gamma}_{P,i}) = \bar{\gamma}_P$*:*

$$\frac{\delta \mathcal{L}(p)}{\delta p}: \quad c\left(\psi - p\right) + \frac{(1-\beta)}{\beta} dn(\bar{\gamma}_P - p) = 0$$

The optimal policy level p, that solves previous equality, is:

$$p_L: \quad p = \frac{c\psi + nd\frac{(1-\beta)}{\beta}\bar{\gamma}_P}{c + nd\frac{(1-\beta)}{\beta}}$$

Whereas transfer will be equal to:

$$t(p_L) = \frac{(1-\beta)}{\beta} d \left[\sum_{i=1}^{n}\left\{(p-p_P)^2 + 2(p-p_P)(p_P - \bar{\gamma}_{P,i})\right\}\right]$$

Proof 3 *By considering* p_L*, it will satisfy lobby participation constraint:*

$$Y - c\left(\psi - p_L\right)^2 - t(p_L) \geq Y - c\left(\psi - p_P\right)^2$$

By substituting the transfer contract defined before, we get:

$$c\left[\left(\psi - p_P\right)^2 - \left(\psi - p_L\right)^2\right] \geq \frac{(1-\beta)}{\beta} d \left[\sum_{i=1}^{n}\left\{(p_L-p_P)^2 + 2(p_L-p_P)(p_P - \bar{\gamma}_{P,i})\right\}\right]$$

It could be rewritten as:

$$c\left[(p_P - p_L)^2 + 2(p_P - p_L)(p - \psi)\right] \geq \frac{(1-\beta)}{\beta} d \left[\sum_{i=1}^{n}\left\{(p_L-p_P)^2 + 2(p_L-p_P)(p_P - \bar{\gamma}_{P,i})\right\}\right]$$

Then, it could be expressed as:

$$\left[c - nd\frac{(1-\beta)}{\beta}\right](p_P - p_L)^2 \geq 2\left[c(p_L - \psi) + \frac{(1-\beta)}{\beta}d\sum_{i=1}^{n}(p_P - \bar{\gamma}_{P,i})\right](p_L - p_P)$$

Since $(p_L - p_P)^2 = (p_P - p_L)^2$*:*

$$\left[c - nd\frac{(1-\beta)}{\beta}\right](p_L - p_P) \geq 2\left[c(p_L - \psi) + \frac{(1-\beta)}{\beta}d\sum_{i=1}^{n}(p_P - \bar{\gamma}_{P,i})\right]$$

$$c\left[(p_L - p_P) - 2(p_L - \psi)\right] \geq nd\frac{(1-\beta)}{\beta}\left[(p_L - p_P) - 2(p_P - \bar{\gamma}_P)\right]$$

$$c\left[(\psi - p_L) + (\psi - p_P)\right] \geq nd\frac{(1-\beta)}{\beta}\left[(p_L - p_P) - 2(p_P - \bar{\gamma}_P)\right]$$

Remember that $p_P = \bar{\gamma}_P$ *and optimal policy under lobby incentive* $p_L = \frac{c\psi + nd\frac{(1-\beta)}{\beta}\bar{\gamma}_P}{c + nd\frac{(1-\beta)}{\beta}}$*:*

$$c\left[\left(\psi - \frac{c\psi + nd\frac{(1-\beta)}{\beta}\bar{\gamma}_P}{c + nd\frac{(1-\beta)}{\beta}}\right) + (\psi - \bar{\gamma}_P)\right] \geq nd\frac{(1-\beta)}{\beta}\left[\frac{c\psi + nd\frac{(1-\beta)}{\beta}\bar{\gamma}_P}{c + nd\frac{(1-\beta)}{\beta}} - \bar{\gamma}_P\right]$$

$$c\left[\left(\frac{nd\frac{(1-\beta)}{\beta}(\psi - \bar{\gamma}_P)}{c + nd\frac{(1-\beta)}{\beta}}\right) + (\psi - \bar{\gamma}_P)\right] \geq nd\frac{(1-\beta)}{\beta}\left[\frac{c(\psi - \bar{\gamma}_P)}{c + nd\frac{(1-\beta)}{\beta}}\right]$$

$$c\left[\left(\frac{nd\frac{(1-\beta)}{\beta}}{c + nd\frac{(1-\beta)}{\beta}}\right) + 1\right] \geq nd\frac{(1-\beta)}{\beta}\left[\frac{c}{c + nd\frac{(1-\beta)}{\beta}}\right]$$

$$c\left[2nd\frac{(1-\beta)}{\beta} + c\right] \geq cnd\frac{(1-\beta)}{\beta}$$

$$nd\frac{(1-\beta)}{\beta} + c \geq 0$$

The pair $\sigma = \left\{p_L, t(p_L)\right\}$ *will be optimal if previous condition is satisfied.*

Individual Trust Simulation

In the simulation we will assume that $n = 701$ individuals set their preference over a policy space depending on the expected state. If the future state is A, their preference distribution will be uniformly distributed between 0 and $\frac{7}{10}$, $\gamma_A \sim Unif[0, \frac{7}{10}]$; whereas if the future state is B, their preference distribution will be uniformly distributed between $\frac{3}{10}$ and 1, $\gamma_B \sim Unif[\frac{1}{4}, 1]$. By implementing a state probability (π) equal to $\frac{1}{2}$, it will generate an expected preference distributed uniformly among $\frac{3}{20}$ and $\frac{17}{20}$, $\bar{\gamma} \sim Unif[\frac{3}{20}, \frac{17}{20}]$.

We need to consider that the individual preference order in the distribution is equal among states, then $\Delta_{\gamma}^{B,A}$ will be equal to $\frac{3}{10}$ for all agents. Furthermore, the individual discounted factor is standardized at $d = 1$.Individuals set their political trust depending on the probability of getting a higher utility with respect to the one commit in electoral competition:

$$E[u_i(p_F)] - E[u_i(p_0)] \geq 0 \Rightarrow (p_F - p_0)\left[(p_F - p_0) + 2(p_0 - \bar{\gamma}_i)\right] \leq 0 \tag{6}$$

Due to timing, individuals generate their expectation with respect to variation in lobby revenue (c), lobby maximizer revenue policy level (ψ) and probability belief distortion (ϵ).In these cases, we will understand that variation in lobby revenue will take an inverse exponential distribution of degree 1, where their density function takes the following shape: $f(c) = \frac{1}{c^2}e^{-\frac{1}{c}}$. Whereas lobby maximizer revenue policy level will be a uniform distribution between 0 and 1 ($\psi \sim Unif[0,1]$) and probability belief distortion will take a uniform distribution in their domain ($\epsilon \sim Unif[\epsilon_{MIN}, \epsilon_{MAX}]$). In order to analyse individual trust in a static framework, we will consider first both pressures separately, and their combination.

Internal Pressure Simulation

In this subsection we will consider that no lobby will suggest a policy to the politician. Then, the unique distortion in their commitment will be a change in the probability belief. If we consider Equation 6, we will observe the following when the implemented policy is p_P

$$(p_P - p_0)\left[(p_P - p_0) + 2(p_0 - \bar{\gamma}_i)\right] \leq 0$$

$$\underbrace{(-\epsilon\Delta_{\gamma}^{B,A})}_{\text{First Condition}}\ \underbrace{\left[-\epsilon\Delta_{\gamma}^{B,A} + 2(\bar{\gamma}_N - \bar{\gamma}_i)\right]}_{\text{Second Condition}} \leq 0$$

Depending on the individual preferred policy position with respect to the averaged one ($\bar{\gamma}_N - \bar{\gamma}_i$), Equation 6 will be satisfied if:

	$(\gamma_N - \gamma_i) \geq 0$	$(\gamma_N - \gamma_i) \leq 0$
First Condition	$\epsilon \geq 0$	$\epsilon \leq 0$
Second Condition	$\frac{2(\gamma_N - \gamma_i)}{\Delta_{\bar{\gamma}}^{B,A}} \geq \epsilon$	$\frac{2(\gamma_N - \gamma_i)}{\Delta_{\bar{\gamma}}^{B,A}} \leq \epsilon$

If we compute the expectation of finding a distortion in probability state into the range, taking into account the domain of the expected variable:

- For $2(\gamma_N - \gamma_i) \geq 0$:

  - When $\frac{2(\gamma_N-\gamma_i)}{\Delta_{\bar{\gamma}}^{B,A}} < \epsilon_{MAX}$:

$$Pr\left[0 \leq \epsilon \leq \frac{2(\gamma_N-\gamma_i)}{\Delta_{\bar{\gamma}}^{B,A}}\right] = \int_0^{\frac{2(\gamma_N-\gamma_i)}{\Delta_{\bar{\gamma}}^{B,A}}} 1 \; d\epsilon$$

  - When $\frac{2(\gamma_N-\gamma_i)}{\Delta_{\bar{\gamma}}^{B,A}} \geq \epsilon_{MAX}$:

$$Pr\left[0 \leq \epsilon \leq \frac{2(\gamma_N-\gamma_i)}{\Delta_{\bar{\gamma}}^{B,A}}\right] = \int_0^{\epsilon_{MAX}} 1 \; d\epsilon$$

- For $2(\gamma_N - \gamma_i) \leq 0$:
  - When $\frac{2(\gamma_N-\gamma_i)}{\Delta_{\bar{\gamma}}^{B,A}} > \epsilon_{MIN}$:

$$Pr\left[\frac{2(\gamma_N-\gamma_i)}{\Delta_{\bar{\gamma}}^{B,A}} \leq \epsilon \leq 0\right] = \int_{\frac{2(\gamma_N-\gamma_i)}{\Delta_{\bar{\gamma}}^{B,A}}}^{0} 1 \; d\epsilon$$

  - When $\frac{2(\gamma_N-\gamma_i)}{\Delta_{\bar{\gamma}}^{B,A}} \leq \epsilon_{MIN}$:

$$Pr\left[\frac{2(\gamma_N-\gamma_i)}{\Delta_{\bar{\gamma}}^{B,A}} \leq \epsilon \leq 0\right] = \int_{\epsilon_{MIN}}^{0} 1 \; d\epsilon$$

The probability of obtaining a higher expected utility will depend on the probability of obtaining a beneficial distortion of probability belief. Furthermore, these ranges must fit with the domain of the variable.

External Pressure Simulation

In this subsection we will consider that no internal pressure is applied in the policy implementation. Then the unique distortion in their commitment will be a potential lobby offer. If we consider Equation 6, we will observe the following when the implemented policy is p^F:

$$(p_P - p_0)\left[(p_P - p_0) + 2(p_0 - \bar{\gamma}_i)\right] \leq 0$$

$$\underbrace{\left(\frac{c(\psi - \bar{\gamma}_N)}{c + nd\frac{(1-\beta)}{\beta}}\right)}_{\text{First Condition}} \underbrace{\left[\left(\frac{c(\psi - \bar{\gamma}_N)}{c + nd\frac{(1-\beta)}{\beta}}\right) + 2(\bar{\gamma}_N - \bar{\gamma}_i)\right]}_{\text{Second Condition}} \leq 0$$

As we have done in previous subsection, depending on the individual preferred policy position with respect to the averaged one $(\bar{\gamma}_N - \bar{\gamma}_i)$, Equation 6 will be satisfied if:

	$(\gamma_N - \gamma_i) \geq 0$	$(\gamma_N - \gamma_i) \leq 0$
First Condition	$\psi \leq \bar{\gamma}_N$	$\psi \geq \bar{\gamma}_N$
Second Condition	$c \leq \frac{2nd\frac{(1-\beta)}{\beta}(\gamma_N-\gamma_i)}{\bar{\gamma}_N-2(\bar{\gamma}_N-\bar{\gamma}_i)-\psi}$	$c \leq \frac{2nd\frac{(1-\beta)}{\beta}(\gamma_N-\gamma_i)}{\bar{\gamma}_N-2(\bar{\gamma}_N-\bar{\gamma}_i)-\psi}$
Extra Condition	$\psi \geq \bar{\gamma}_N - 2(\bar{\gamma}_N - \bar{\gamma}_i)$	$\psi \leq \bar{\gamma}_N - 2(\bar{\gamma}_N - \bar{\gamma}_i)$

Starting by *First Condition*, it is straightforward to observe that in order to improve individual utility lobby maximizer revenue policy must be located close to the individual most preferred policy. With respect to the *Second Condition*, this restriction helps to keep the final implemented policy in the improving region. For the existence of this restriction, we need a positive argument in the cost limit, generated by *Extra Condition.* Then, if we compute the likelihood of finding a lobby characteristics into the range, we observe:

- For $2(\gamma_N - \gamma_i) \geq 0$:
    - When $\bar{\gamma}_N - 2(\bar{\gamma}_N - \bar{\gamma}_i) > 0$:

$$Pr\left[\left\{\bar{\gamma}_N - 2(\bar{\gamma}_N - \bar{\gamma}_i) \leq \psi \leq \bar{\gamma}_N\right\} \cap \left\{0 \leq c \leq \frac{2nd\frac{(1-\beta)}{\beta}(\gamma_N - \gamma_i)}{\bar{\gamma}_N - 2(\bar{\gamma}_N - \bar{\gamma}_i) - \psi}\right\}\right]$$

$$= \int_{\bar{\gamma}_N-2(\bar{\gamma}_N-\bar{\gamma}_i)}^{\bar{\gamma}_N} 1\ d\psi + \int_0^{\bar{\gamma}_N-2(\bar{\gamma}_N-\bar{\gamma}_i)} \int_0^{\frac{2nd\frac{(1-\beta)}{\beta}(\gamma_N-\gamma_i)}{\bar{\gamma}_N-2(\bar{\gamma}_N-\bar{\gamma}_i)-\psi}} \frac{1}{c^2}e^{-\frac{1}{c}}\ dc\ d\psi$$

    - When $\bar{\gamma}_N - 2(\bar{\gamma}_N - \bar{\gamma}_i) \leq 0$:

$$Pr\left[\left\{\bar{\gamma}_N - 2(\bar{\gamma}_N - \bar{\gamma}_i) \leq \psi \leq \bar{\gamma}_N\right\} \cap \left\{0 \leq c \leq \infty\right\}\right] = \int_0^{\bar{\gamma}_N} 1\ d\psi$$

- For $2(\gamma_N - \gamma_i) \leq 0$:
    - When $\bar{\gamma}_N - 2(\bar{\gamma}_N - \bar{\gamma}_i) < 1$:

$$Pr\left[\left\{\bar{\gamma}_N \leq \psi \leq \bar{\gamma}_N - 2(\bar{\gamma}_N - \bar{\gamma}_i)\right\} \cap \left\{0 \leq c \leq \frac{2nd\frac{(1-\beta)}{\beta}(\gamma_N - \gamma_i)}{\bar{\gamma}_N - 2(\bar{\gamma}_N - \bar{\gamma}_i) - \psi}\right\}\right]$$

$$= \int_{\bar{\gamma}_N}^{\bar{\gamma}_N-2(\bar{\gamma}_N-\bar{\gamma}_i)} 1\ d\psi + \int_{bar\gamma_N-2(\bar{\gamma}_N-\bar{\gamma}_i)}^{1} \int_0^{\frac{2nd\frac{(1-\beta)}{\beta}(\gamma_N-\gamma_i)}{\bar{\gamma}_N-2(\bar{\gamma}_N-\bar{\gamma}_i)-\psi}} \frac{1}{c^2}e^{-\frac{1}{c}}\ dc\ d\psi$$

    - When $\bar{\gamma}_N \leq \bar{\gamma}_N - 2(\bar{\gamma}_N - \bar{\gamma}_i)$:

$$Pr\left[\left\{\bar{\gamma}_N \leq \psi \leq \bar{\gamma}_N - 2(\bar{\gamma}_N - \bar{\gamma}_i)\right\} \cap \left\{0 \leq c \leq \infty\right\}\right] = \int_{\bar{\gamma}_N}^{1} 1\ d\psi$$

As we have computed in previous subsection, the probability of obtaining a higher expected utility will depend on the expectation with respect to the lobby characteristics.

Internal and External Pressure Simulation

In this subsection, we will analyse the behaviour of individual trust when there is a distortion in the state-probability belief and under the existence of a lobby. At this point, the expectation of the lobby characteristics will determine the level of individual trust. Considering p_F as the implemented final policy under state-probability distortion and the existence of a lobby, Equation 6 will be satisfied if:

$$\underbrace{\left(\frac{c(\psi-\bar{\gamma}_N)+nd\frac{(1-\beta)}{\beta}\left(\epsilon\Delta_{\gamma}^{B,A}\right)}{c+nd\frac{(1-\beta)}{\beta}}\right)}_{\text{First Condition}}\underbrace{\left[\left(\frac{c(\psi-\bar{\gamma}_N)+nd\frac{(1-\beta)}{\beta}\left(\epsilon\Delta_{\gamma}^{B,A}\right)}{c+nd\frac{(1-\beta)}{\beta}}\right)+2(\bar{\gamma}_N-\bar{\gamma}_i)\right]}_{\text{Second Condition}}\leq 0$$

Depending on the individual preferred policy with respect to the averaged one $(\bar{\gamma}_N-\bar{\gamma}_i)$, Equation 6 will be satisfied if:

	$(\bar{\gamma}_N-\bar{\gamma}_i)\geq 0$	$(\bar{\gamma}_N-\bar{\gamma}_i)\leq 0$
First Condition	$c\geq\frac{nd\frac{(1-\beta)}{\beta}\epsilon\Delta_{\gamma}^{B,A}}{(\bar{\gamma}_N-\psi)}$	$c\leq\frac{nd\frac{(1-\beta)}{\beta}\epsilon\Delta_{\gamma}^{B,A}}{(\bar{\gamma}_N-\psi)}$
Second Condition	$c\leq\frac{nd\frac{(1-\beta)}{\beta}\left[\Delta_{\gamma}^{B,A}\epsilon+2(\gamma_N-\gamma_i)\right]}{\bar{\gamma}_N-2(\bar{\gamma}_N-\bar{\gamma}_i)-\psi}$	$c\geq\frac{nd\frac{(1-\beta)}{\beta}\left[\Delta_{\gamma}^{B,A}\epsilon+2(\gamma_N-\gamma_i)\right]}{\bar{\gamma}_N-2(\bar{\gamma}_N-\bar{\gamma}_i)-\psi}$

In this case, depending on the individual most preferred policy position, their probability belief will be characterized by:

a) When $\epsilon_{MIN}\geq\frac{-2(\bar{\gamma}_N-\bar{\gamma}_i)}{\Delta_{\gamma}^{B,A}}$ and:

  a.i) $0\geq\bar{\gamma}_N-2(\bar{\gamma}_N-\bar{\gamma}_i)$, individual will compute their individual trust through the computation of following probabilities:

    I) When state-probability distortion and lobby maximizer revenue policy weakly increases the individual utility function that it will occur if $\epsilon\in[\epsilon_{MIN},0]$ and $\psi\in[0,\bar{\gamma}]$, respectively. Then, probability of this scenario is:

$$\Pr\left(\{\epsilon\in[\epsilon_{MIN},0]\}\cap\{\psi\in[0,\bar{\gamma}]\}\right)=\int_{\epsilon_{MIN}}^{0}\int_{0}^{\bar{\gamma}_N}1\;d\psi\;d\epsilon$$

    Notice that this situation does not require of cost bounds since both parameters will increase individual utility function regardless the cost level.

    II) When there is a beneficial state-probability distortion ($\epsilon\in[\epsilon_{MIN},0]$), but a disadvantageous lobby maximizer revenue policy ($\psi\in[\bar{\gamma},1]$). This situation will improve individual utility function if $c\in\left[0,\frac{nd\frac{(1-\beta)}{\beta}\epsilon\Delta_{\gamma}^{B,A}}{(\bar{\gamma}_N-\psi)}\right]$.

Then, probability of this scenario is:

$$\Pr\left(\{\epsilon \in [\epsilon_{MIN}, 0]\} \cap \{\psi \in [0, \bar{\gamma}]\} \cap \left\{c \in \left[0, \frac{nd\frac{(1-\beta)}{\beta}\epsilon\Delta_{\gamma}^{B,A}}{(\bar{\gamma}_N - \psi)}\right]\right\}\right)$$

$$= \int_{\epsilon_{MIN}}^{0} \int_{0}^{\bar{\gamma}_N} \int_{0}^{\frac{nd\frac{(1-\beta)}{\beta}\epsilon\Delta_{\gamma}^{B,A}}{(\bar{\gamma}_N-\psi)}} \frac{1}{c^2} e^{-\frac{1}{c}} \ dc \ d\psi \ d\epsilon$$

Given this scenario, lobby revenue variation located in the specified region will improve individual utility function.

III) When there is a beneficial lobby maximizer revenue policy ($\psi \in [0, \bar{\gamma}_N]$), but a disadvantageous state-probability distortion ($\epsilon \in [0, \epsilon_{MAX}]$). This situation will improve individual utility function if $c \in \left[\frac{nd\frac{(1-\beta)}{\beta}\epsilon\Delta_{\gamma}^{B,A}}{(\bar{\gamma}_N-\psi)}, \infty\right]$. Then, probability of this scenario is:

$$\Pr\left(\{\epsilon \in [0, \epsilon_{MAX}]\} \cap \{\psi \in [\bar{\gamma}_N, 1]\} \cap \left\{c \in \left[\frac{nd\frac{(1-\beta)}{\beta}\epsilon\Delta_{\gamma}^{B,A}}{(\bar{\gamma}_N - \psi)}, \infty\right]\right\}\right)$$

$$= \int_{0}^{\epsilon_{MAX}} \int_{\bar{\gamma}_N}^{1} \int_{\frac{nd\frac{(1-\beta)}{\beta}\epsilon\Delta_{\gamma}^{B,A}}{(\bar{\gamma}_N-\psi)}}^{\infty} \frac{1}{c^2} e^{-\frac{1}{c}} \ dc \ d\psi \ d\epsilon$$

Given this scenario, lobby revenue variation located in the specified region will improve individual utility function.

Then, for those individuals that are characterized for this most preferred policy, their individual trust function will be computed as:

$$\mathbb{T}_N(\bar{\gamma}_i) = \sum_{j=I}^{III} Pr(\epsilon, \psi, c)$$

In other words, by aggregating three computed probabilities, we will get the individual trust function of an individual characterized by their most preferred policy $\bar{\gamma}_i$.

a.ii) $0 < \bar{\gamma}_N - 2(\bar{\gamma}_N - \bar{\gamma}_i) \leq \bar{\gamma}_N$, individual will compute their individual trust through the computation of following probabilities:

I) When state-probability distortion and lobby maximizer revenue policy weakly increases the individual utility function that it will occur if $\epsilon \in [\epsilon_{MIN}, 0]$ and $\psi \in [\bar{\gamma}_N - 2(\bar{\gamma}_N - \bar{\gamma}_i), \bar{\gamma}_N]$, respectively. Then, probability of this scenario is:

$$\Pr\left(\{\epsilon \in [\epsilon_{MIN}, 0]\} \cap \{\psi \in [\bar{\gamma}_N - 2(\bar{\gamma}_N - \bar{\gamma}_i), \bar{\gamma}]\}\right) = \int_{\epsilon_{MIN}}^{0} \int_{\bar{\gamma}_N - 2(\bar{\gamma}_N - \bar{\gamma}_i)}^{\bar{\gamma}_N} 1 \ d\psi \ d\epsilon$$

Notice that this situation does not require of cost bounds since both parameters will increase individual utility function regardless the cost level.

II) When there is a beneficial state-probability distortion ($\epsilon \in [\epsilon_{MIN}, 0]$), but a disadvantageous lobby maximizer revenue policy to the right side($\psi \in [\bar{\gamma}_N, 1]$). This situation will improve individual utility function if $c \in \left[0, \frac{nd\frac{(1-\beta)}{\beta}\epsilon\Delta_{\gamma}^{B,A}}{(\bar{\gamma}_N - \psi)}\right]$. Then, probability of this scenario is:

$$\Pr\left(\{\epsilon \in [\epsilon_{MIN}, 0]\} \cap \{\psi \in [\bar{\gamma}_N, 1]\} \cap \left\{c \in \left[0, \frac{nd\frac{(1-\beta)}{\beta}\epsilon\Delta_{\gamma}^{B,A}}{(\bar{\gamma}_N - \psi)}\right]\right\}\right)$$

$$= \int_{\epsilon_{MIN}}^{0} \int_{\bar{\gamma}_N}^{1} \int_{0}^{\frac{nd\frac{(1-\beta)}{\beta}\epsilon\Delta_{\gamma}^{B,A}}{(\bar{\gamma}_N - \psi)}} \frac{1}{c^2} e^{-\frac{1}{c}}\, dc\, d\psi\, d\epsilon$$

Given this scenario, lobby revenue variation located in the specified region will improve individual utility function.

III) When there is a beneficial state-probability distortion ($\epsilon \in [\epsilon_{MIN}, 0]$), but a disadvantageous lobby maximizer revenue policy to the left side ($\psi \in [0, \bar{\gamma}_N - 2(\bar{\gamma}_N - \bar{\gamma}_i)]$). This situation will improve individual utility function if $c \in \left[0, \frac{nd\frac{(1-\beta)}{\beta}\left[\Delta_{\gamma}^{B,A}\epsilon + 2(\gamma_N - \gamma_i)\right]}{\bar{\gamma}_N - 2(\bar{\gamma}_N - \bar{\gamma}_i) - \psi}\right]$. Then, probability of this scenario is:

$$\Pr\left[\left\{\epsilon \in [\epsilon_{MIN}, 0]\right\} \cap \left\{\psi \in [0, \bar{\gamma}_N - 2(\bar{\gamma}_N - \bar{\gamma}_i)]\right\} \cap ...\right.$$

$$\left. ... \cap \left\{c \in \left[0, \frac{nd\frac{(1-\beta)}{\beta}\left[\Delta_{\gamma}^{B,A}\epsilon + 2(\gamma_N - \gamma_i)\right]}{\bar{\gamma}_N - 2(\bar{\gamma}_N - \bar{\gamma}_i) - \psi}\right]\right\}\right] =$$

$$= \int_{\epsilon_{MIN}}^{0} \int_{0}^{\bar{\gamma}_N - 2(\bar{\gamma}_N - \bar{\gamma}_i)} \int_{0}^{\frac{nd\frac{(1-\beta)}{\beta}\left[\Delta_{\gamma}^{B,A}\epsilon + 2(\gamma_N - \gamma_i)\right]}{\bar{\gamma}_N - 2(\bar{\gamma}_N - \bar{\gamma}_i) - \psi}} \frac{1}{c^2} e^{-\frac{1}{c}}\, dc\, d\psi\, d\epsilon$$

Given this scenario, lobby revenue variation located in the specified region will improve individual utility function.

IV) When there is a beneficial lobby maximizer revenue policy ($\psi \in [\bar{\gamma}_N - 2(\bar{\gamma}_N - \bar{\gamma}_i), \bar{\gamma}_N]$), but a disadvantageous state-probability distortion to the right side ($\epsilon \in [0, \epsilon_{MAX}]$). This situation will improve individual utility function if $c \in \left[\frac{nd\frac{(1-\beta)}{\beta}\epsilon\Delta_{\gamma}^{B,A}}{(\bar{\gamma}_N - \psi)}, \infty\right]$. Then, probability of this scenario is:

$$\Pr\left(\{\epsilon \in [0, \epsilon_{MAX}]\} \cap \{\psi \in [\bar{\gamma}_N - 2(\bar{\gamma}_N - \bar{\gamma}_i), \bar{\gamma}_N]\} \cap \left\{c \in \left[\frac{nd\frac{(1-\beta)}{\beta}\epsilon\Delta_{\gamma}^{B,A}}{(\bar{\gamma}_N - \psi)}, \infty\right]\right\}\right)$$

$$= \int_{0}^{\epsilon_{MAX}} \int_{\bar{\gamma}_N - 2(\bar{\gamma}_N - \bar{\gamma}_i)}^{\bar{\gamma}_N} \int_{\frac{nd\frac{(1-\beta)}{\beta}\epsilon\Delta_{\gamma}^{B,A}}{(\bar{\gamma}_N - \psi)}}^{\infty} \frac{1}{c^2} e^{-\frac{1}{c}} \; dc \; d\psi \; d\epsilon$$

Given this scenario, lobby revenue variation located in the specified region will improve individual utility function.

V) When both state-probability distortion (to the right side) and lobby maximizer revenue policy (to the left side) are disadvantageous ($\epsilon \in [0, \epsilon_{MAX}]$ and $\psi \in [0, \bar{\gamma}_N - 2(\bar{\gamma}_N - \bar{\gamma}_i)]$, respectively), but a certain level of lobby revenue variation weakly increases the individual utility function, $c \in \left[\frac{nd\frac{(1-\beta)}{\beta}\epsilon\Delta_{\gamma}^{B,A}}{(\bar{\gamma}_N - \psi)}, \frac{nd\frac{(1-\beta)}{\beta}\left[\Delta_{\gamma}^{B,A}\epsilon + 2(\gamma_N - \gamma_i)\right]}{\bar{\gamma}_N - 2(\bar{\gamma}_N - \bar{\gamma}_i) - \psi}\right]$. Then, probability of this scenario is:

$$\Pr\left[\left\{\epsilon \in [0, \epsilon_{MAX}]\right\} \cap \left\{\psi \in [0, \bar{\gamma}_N - 2(\bar{\gamma}_N - \bar{\gamma}_i)]\right\} \cap ...\right.$$

$$\left. ... \cap \left\{c \in \left[\frac{nd\frac{(1-\beta)}{\beta}\epsilon\Delta_{\gamma}^{B,A}}{(\bar{\gamma}_N - \psi)}, \frac{nd\frac{(1-\beta)}{\beta}\left[\Delta_{\gamma}^{B,A}\epsilon + 2(\gamma_N - \gamma_i)\right]}{\bar{\gamma}_N - 2(\bar{\gamma}_N - \bar{\gamma}_i) - \psi}\right]\right\}\right] =$$

$$= \int_{0}^{\epsilon_{MAX}} \int_{0}^{\bar{\gamma}_N - 2(\bar{\gamma}_N - \bar{\gamma}_i)} \int_{\frac{nd\frac{(1-\beta)}{\beta}\epsilon\Delta_{\gamma}^{B,A}}{(\bar{\gamma}_N - \psi)}}^{\frac{nd\frac{(1-\beta)}{\beta}\left[\Delta_{\gamma}^{B,A}\epsilon + 2(\gamma_N - \gamma_i)\right]}{\bar{\gamma}_N - 2(\bar{\gamma}_N - \bar{\gamma}_i) - \psi}} \frac{1}{c^2} e^{-\frac{1}{c}} \; dc \; d\psi \; d\epsilon$$

Given this scenario, lobby revenue variation located in the specified region will improve individual utility function.

Then, for those individuals that are characterized for this most preferred policy, their individual trust function will be computed as:

$$\mathbb{T}_N(\bar{\gamma}_i) = \sum_{j=I}^{V} Pr(\epsilon, \psi, c)$$

b) When $\epsilon_{MIN} < \frac{-2(\bar{\gamma}_N - \bar{\gamma}_i)}{\Delta_{\gamma}^{B,A}}$ and $0 < \bar{\gamma}_N - 2(\bar{\gamma}_N - \bar{\gamma}_i) \leq \bar{\gamma}_N$, individual will compute their individual trust through the computation of following probabilities:

I) When state-probability distortion and lobby maximizer revenue policy weakly increases the individual utility function that it will occur if $\epsilon \in [\frac{-2(\bar{\gamma}_N - \bar{\gamma}_i)}{\Delta_{\gamma}^{B,A}}, 0]$ and $\psi \in [\bar{\gamma}_N - 2(\bar{\gamma}_N - \bar{\gamma}_i), \bar{\gamma}_N]$, respectively. Then, probability of this scenario is:

$$\Pr\left(\left\{\epsilon \in [\frac{-2(\bar{\gamma}_N - \bar{\gamma}_i)}{\Delta_{\gamma}^{B,A}}, 0]\right\} \cap \{\psi \in [\bar{\gamma}_N - 2(\bar{\gamma}_N - \bar{\gamma}_i), \bar{\gamma}]\}\right)$$

$$= \int_{\frac{-2(\bar{\gamma}_N - \bar{\gamma}_i)}{\Delta_{\gamma}^{B,A}}}^{0} \int_{\bar{\gamma}_N - 2(\bar{\gamma}_N - \bar{\gamma}_i)}^{\bar{\gamma}_N} 1 \; d\psi \; d\epsilon$$

Notice that this situation does not require of cost bounds since both parameters will increase individual utility function regardless the cost level.

II) When there is a beneficial state-probability distortion ($\epsilon \in [\frac{-2(\bar{\gamma}_N - \bar{\gamma}_i)}{\Delta_\gamma^{B,A}}, 0]$), but a disadvantageous lobby maximizer revenue policy to the right side($\psi \in [\bar{\gamma}_N, 1]$). This situation will improve individual utility function if $c \in \left[0, \frac{nd\frac{(1-\beta)}{\beta}\epsilon\Delta_\gamma^{B,A}}{(\bar{\gamma}_N - \psi)}\right]$. Then, probability of this scenario is:

$$\Pr\left(\left\{\epsilon \in \left[\frac{-2(\bar{\gamma}_N - \bar{\gamma}_i)}{\Delta_\gamma^{B,A}}, 0\right]\right\} \cap \{\psi \in [\bar{\gamma}_N, 1]\} \cap \left\{c \in \left[0, \frac{nd\frac{(1-\beta)}{\beta}\epsilon\Delta_\gamma^{B,A}}{(\bar{\gamma}_N - \psi)}\right]\right\}\right)$$

$$= \int_{\frac{-2(\bar{\gamma}_N - \bar{\gamma}_i)}{\Delta_\gamma^{B,A}}}^{0} \int_{\bar{\gamma}_N}^{1} \int_{0}^{\frac{nd\frac{(1-\beta)}{\beta}\epsilon\Delta_\gamma^{B,A}}{(\bar{\gamma}_N - \psi)}} \frac{1}{c^2} e^{-\frac{1}{c}} \, dc \, d\psi \, d\epsilon$$

Given this scenario, lobby revenue variation located in the specified region will improve individual utility function.

III) When there is a beneficial state-probability distortion ($\epsilon \in [\frac{-2(\bar{\gamma}_N - \bar{\gamma}_i)}{\Delta_\gamma^{B,A}}, 0]$), but a disadvantageous lobby maximizer revenue policy to the left side ($\psi \in [0, \bar{\gamma}_N - 2(\bar{\gamma}_N - \bar{\gamma}_i)]$). This situation will improve individual utility function if $c \in \left[0, \frac{nd\frac{(1-\beta)}{\beta}\left[\Delta_\gamma^{B,A}\epsilon + 2(\gamma_N - \gamma_i)\right]}{\bar{\gamma}_N - 2(\bar{\gamma}_N - \bar{\gamma}_i) - \psi}\right]$. Then, probability of this scenario is:

$$\Pr\left[\left\{\epsilon \in [\frac{-2(\bar{\gamma}_N - \bar{\gamma}_i)}{\Delta_\gamma^{B,A}}, 0]\right\} \cap \left\{\psi \in [0, \bar{\gamma}_N - 2(\bar{\gamma}_N - \bar{\gamma}_i)]\right\} \cap ...\right.$$

$$\left. ... \cap \left\{c \in \left[0, \frac{nd\frac{(1-\beta)}{\beta}\left[\Delta_\gamma^{B,A}\epsilon + 2(\gamma_N - \gamma_i)\right]}{\bar{\gamma}_N - 2(\bar{\gamma}_N - \bar{\gamma}_i) - \psi}\right]\right\}\right] =$$

$$= \int_{\frac{-2(\bar{\gamma}_N - \bar{\gamma}_i)}{\Delta_\gamma^{B,A}}}^{0} \int_{0}^{\bar{\gamma}_N - 2(\bar{\gamma}_N - \bar{\gamma}_i)} \int_{0}^{\frac{nd\frac{(1-\beta)}{\beta}\left[\Delta_\gamma^{B,A}\epsilon + 2(\gamma_N - \gamma_i)\right]}{\bar{\gamma}_N - 2(\bar{\gamma}_N - \bar{\gamma}_i) - \psi}} \frac{1}{c^2} e^{-\frac{1}{c}} \, dc \, d\psi \, d\epsilon$$

Given this scenario, lobby revenue variation located in the specified region will improve individual utility function.

IV) When there is a beneficial lobby maximizer revenue policy ($\psi \in [\bar{\gamma}_N - 2(\bar{\gamma}_N - \bar{\gamma}_i), \bar{\gamma}_N]$), but a disadvantageous state-probability distortion to the right side ($\epsilon \in [0, \epsilon_{MAX}]$). This situation will improve individual utility function if $c \in \left[\frac{nd\frac{(1-\beta)}{\beta}\epsilon\Delta_\gamma^{B,A}}{(\bar{\gamma}_N - \psi)}, \infty\right]$. Then, probability of this scenario is:

$$\Pr\left(\{\epsilon \in [0, \epsilon_{MAX}]\} \cap \{\psi \in [\bar{\gamma}_N - 2(\bar{\gamma}_N - \bar{\gamma}_i), \bar{\gamma}_N]\} \cap \left\{c \in \left[\frac{nd\frac{(1-\beta)}{\beta}\epsilon\Delta_\gamma^{B,A}}{(\bar{\gamma}_N - \psi)}, \infty\right]\right\}\right)$$

$$= \int_{0}^{\epsilon_{MAX}} \int_{\bar{\gamma}_N - 2(\bar{\gamma}_N - \bar{\gamma}_i)}^{\bar{\gamma}_N} \int_{\frac{nd\frac{(1-\beta)}{\beta}\epsilon\Delta_{\gamma}^{B,A}}{(\bar{\gamma}_N - \psi)}}^{\infty} \frac{1}{c^2} e^{-\frac{1}{c}} \ dc \ d\psi \ d\epsilon$$

Given this scenario, lobby revenue variation located in the specified region will improve individual utility function.

V) When there is a beneficial lobby maximizer revenue policy ($\psi \in [\bar{\gamma}_N - 2(\bar{\gamma}_N - \bar{\gamma}_i), \bar{\gamma}_N]$), but a disadvantageous state-probability distortion to the left side ($\epsilon \in [\epsilon_{MIN}, \frac{-2(\bar{\gamma}_N - \bar{\gamma}_i)}{\Delta_{\gamma}^{B,A}}]$). This situation will improve individual utility function if $c \in \left[\frac{nd\frac{(1-\beta)}{\beta}\left[\Delta_{\gamma}^{B,A}\epsilon + 2(\gamma_N - \gamma_i)\right]}{\bar{\gamma}_N - 2(\bar{\gamma}_N - \bar{\gamma}_i) - \psi}, \infty\right]$. Then, probability of this scenario is:

$$\Pr\left[\left\{\epsilon \in [\epsilon_{MIN}, \frac{-2(\bar{\gamma}_N - \bar{\gamma}_i)}{\Delta_{\gamma}^{B,A}}]\right\} \cap \left\{\psi \in [\bar{\gamma}_N - 2(\bar{\gamma}_N - \bar{\gamma}_i), \bar{\gamma}_N]\right\} \cap ...\right.$$

$$\left. ... \cap \left\{c \in \left[\frac{nd\frac{(1-\beta)}{\beta}\left[\Delta_{\gamma}^{B,A}\epsilon + 2(\gamma_N - \gamma_i)\right]}{\bar{\gamma}_N - 2(\bar{\gamma}_N - \bar{\gamma}_i) - \psi}, \infty\right]\right\}\right] =$$

$$= \int_{\epsilon_{MIN}}^{\frac{-2(\bar{\gamma}_N - \bar{\gamma}_i)}{\Delta_{\gamma}^{B,A}}} \int_{\bar{\gamma}_N - 2(\bar{\gamma}_N - \bar{\gamma}_i)}^{\bar{\gamma}_N} \int_{\frac{nd\frac{(1-\beta)}{\beta}\left[\Delta_{\gamma}^{B,A}\epsilon + 2(\gamma_N - \gamma_i)\right]}{\bar{\gamma}_N - 2(\bar{\gamma}_N - \bar{\gamma}_i) - \psi}}^{\infty} \frac{1}{c^2} e^{-\frac{1}{c}} \ dc \ d\psi \ d\epsilon$$

Given this scenario, lobby revenue variation located in the specified region will improve individual utility function.

VI) When both state-probability distortion (to the right side) and lobby maximizer revenue policy (to the left side) are disadvantageous ($\epsilon \in [0, \epsilon_{MAX}]$ and $\psi \in [0, \bar{\gamma}_N - 2(\bar{\gamma}_N - \bar{\gamma}_i)]$, respectively), but a certain level of lobby revenue variation weakly increases the individual utility function, $c \in \left[\frac{nd\frac{(1-\beta)}{\beta}\epsilon\Delta_{\gamma}^{B,A}}{(\bar{\gamma}_N - \psi)}, \frac{nd\frac{(1-\beta)}{\beta}\left[\Delta_{\gamma}^{B,A}\epsilon + 2(\gamma_N - \gamma_i)\right]}{\bar{\gamma}_N - 2(\bar{\gamma}_N - \bar{\gamma}_i) - \psi}\right]$. Then, probability of this scenario is:

$$\Pr\left[\left\{\epsilon \in [0, \epsilon_{MAX}]\right\} \cap \left\{\psi \in [0, \bar{\gamma}_N - 2(\bar{\gamma}_N - \bar{\gamma}_i)]\right\} \cap ...\right.$$

$$\left. ... \cap \left\{c \in \left[\frac{nd\frac{(1-\beta)}{\beta}\epsilon\Delta_{\gamma}^{B,A}}{(\bar{\gamma}_N - \psi)}, \frac{nd\frac{(1-\beta)}{\beta}\left[\Delta_{\gamma}^{B,A}\epsilon + 2(\gamma_N - \gamma_i)\right]}{\bar{\gamma}_N - 2(\bar{\gamma}_N - \bar{\gamma}_i) - \psi}\right]\right\}\right] =$$

$$= \int_{0}^{\epsilon_{MAX}} \int_{0}^{\bar{\gamma}_N - 2(\bar{\gamma}_N - \bar{\gamma}_i)} \int_{\frac{nd\frac{(1-\beta)}{\beta}\epsilon\Delta_{\gamma}^{B,A}}{(\bar{\gamma}_N - \psi)}}^{\frac{nd\frac{(1-\beta)}{\beta}\left[\Delta_{\gamma}^{B,A}\epsilon + 2(\gamma_N - \gamma_i)\right]}{\bar{\gamma}_N - 2(\bar{\gamma}_N - \bar{\gamma}_i) - \psi}} \frac{1}{c^2} e^{-\frac{1}{c}} \ dc \ d\psi \ d\epsilon$$

Given this scenario, lobby revenue variation located in the specified region will improve individual utility function.

VII) When both state-probability distortion (to the left side) and lobby maximizer revenue policy (to the right side) are disadvantageous ($\epsilon \in$

$[\epsilon_{MIN}, \frac{-2(\bar{\gamma}_N-\bar{\gamma}_i)}{\Delta_\gamma^{B,A}}]$ and $\psi \in [\bar{\gamma}_N, 1]$, respectively), but a certain level of lobby revenue variation weakly increases the individual utility function, $c \in \left[\frac{nd\frac{(1-\beta)}{\beta}\left[\Delta_\gamma^{B,A}\epsilon+2(\gamma_N-\gamma_i)\right]}{\bar{\gamma}_N-2(\bar{\gamma}_N-\bar{\gamma}_i)-\psi}, \frac{nd\frac{(1-\beta)}{\beta}\epsilon\Delta_\gamma^{B,A}}{(\bar{\gamma}_N-\psi)}\right]$. Then, probability of this scenario is:

$$\Pr\left[\left\{\epsilon \in [\epsilon_{MIN}, \frac{-2(\bar{\gamma}_N-\bar{\gamma}_i)}{\Delta_\gamma^{B,A}}]\right\} \cap \left\{\psi \in [\bar{\gamma}_N, 1]\right\} \cap ...\right.$$

$$\left.... \cap \left\{c \in \left[\frac{nd\frac{(1-\beta)}{\beta}\left[\Delta_\gamma^{B,A}\epsilon+2(\gamma_N-\gamma_i)\right]}{\bar{\gamma}_N-2(\bar{\gamma}_N-\bar{\gamma}_i)-\psi}, \frac{nd\frac{(1-\beta)}{\beta}\epsilon\Delta_\gamma^{B,A}}{(\bar{\gamma}_N-\psi)}\right]\right\}\right] =$$

$$= \int_{\epsilon_{MIN}}^{\frac{-2(\bar{\gamma}_N-\bar{\gamma}_i)}{\Delta_\gamma^{B,A}}} \int_{\bar{\gamma}_N}^{1} \int_{\frac{nd\frac{(1-\beta)}{\beta}\left[\Delta_\gamma^{B,A}\epsilon+2(\gamma_N-\gamma_i)\right]}{\bar{\gamma}_N-2(\bar{\gamma}_N-\bar{\gamma}_i)-\psi}}^{\frac{nd\frac{(1-\beta)}{\beta}\epsilon\Delta_\gamma^{B,A}}{(\bar{\gamma}_N-\psi)}} \frac{1}{c^2}e^{-\frac{1}{c}}\ dc\ d\psi\ d\epsilon$$

Given this scenario, lobby revenue variation located in the specified region will improve individual utility function.

Then, for those individuals that are characterized for this most preferred policy, their individual trust function will be computed as:

$$\mathbb{T}_N(\bar{\gamma}_i) = \sum_{j=I}^{VII} Pr(\epsilon, \psi, c)$$

c) When $\epsilon_{MAX} > \frac{-2(\bar{\gamma}_N-\bar{\gamma}_i)}{\Delta_\gamma^{B,A}}$ and $\bar{\gamma}_N \le \bar{\gamma}_N - 2(\bar{\gamma}_N - \bar{\gamma}_i) < 1$, individual will compute their individual trust through the computation of following probabilities:

I) When state-probability distortion and lobby maximizer revenue policy weakly increases the individual utility function that it will occur if $\epsilon \in [0, \frac{-2(\bar{\gamma}_N-\bar{\gamma}_i)}{\Delta_\gamma^{B,A}}]$ and $\psi \in [\bar{\gamma}_N, \bar{\gamma}_N - 2(\bar{\gamma}_N - \bar{\gamma}_i)]$, respectively. Then, probability of this scenario is:

$$\Pr\left(\left\{\epsilon \in [0, \frac{-2(\bar{\gamma}_N-\bar{\gamma}_i)}{\Delta_\gamma^{B,A}}]\right\} \cap \{\bar{\gamma}_N, \psi \in [\bar{\gamma}_N - 2(\bar{\gamma}_N - \bar{\gamma}_i)]\}\right)$$

$$= \int_{0}^{\frac{-2(\bar{\gamma}_N-\bar{\gamma}_i)}{\Delta_\gamma^{B,A}}} \int_{\bar{\gamma}_N}^{\bar{\gamma}_N-2(\bar{\gamma}_N-\bar{\gamma}_i)} 1\ d\psi\ d\epsilon$$

Notice that this situation does not require of cost bounds since both parameters will increase individual utility function regardless the cost level.

II) When there is a beneficial state-probability distortion ($\epsilon \in [0, \frac{-2(\bar{\gamma}_N-\bar{\gamma}_i)}{\Delta_\gamma^{B,A}}]$), but a disadvantageous lobby maximizer revenue policy to the left side ($\psi \in [0, \bar{\gamma}_N]$). This situation will improve individual utility function if $c \in \left[0, \frac{nd\frac{(1-\beta)}{\beta}\epsilon\Delta_\gamma^{B,A}}{(\bar{\gamma}_N-\psi)}\right]$.

Then, probability of this scenario is:

$$\Pr\left(\left\{\epsilon \in \left[0, \frac{-2(\bar{\gamma}_N - \bar{\gamma}_i)}{\Delta_\gamma^{B,A}}\right]\right\} \cap \{\psi \in [0, \bar{\gamma}_N]\} \cap \left\{c \in \left[0, \frac{nd\frac{(1-\beta)}{\beta}\epsilon\Delta_\gamma^{B,A}}{(\bar{\gamma}_N - \psi)}\right]\right\}\right)$$

$$= \int_0^{\frac{-2(\bar{\gamma}_N - \bar{\gamma}_i)}{\Delta_\gamma^{B,A}}} \int_0^{\bar{\gamma}_N} \int_0^{\frac{nd\frac{(1-\beta)}{\beta}\epsilon\Delta_\gamma^{B,A}}{(\bar{\gamma}_N - \psi)}} \frac{1}{c^2} e^{-\frac{1}{c}} \, dc \, d\psi \, d\epsilon$$

Given this scenario, lobby revenue variation located in the specified region will improve individual utility function.

III) When there is a beneficial state-probability distortion ($\epsilon \in [0, \frac{-2(\bar{\gamma}_N - \bar{\gamma}_i)}{\Delta_\gamma^{B,A}}]$), but a disadvantageous lobby maximizer revenue policy to the right side ($\psi \in [\bar{\gamma}_N - 2(\bar{\gamma}_N - \bar{\gamma}_i), 1]$). This situation will improve individual utility function if $c \in \left[0, \frac{nd\frac{(1-\beta)}{\beta}\left[\Delta_\gamma^{B,A}\epsilon + 2(\gamma_N - \gamma_i)\right]}{\bar{\gamma}_N - 2(\bar{\gamma}_N - \bar{\gamma}_i) - \psi}\right]$. Then, probability of this scenario is:

$$\Pr\left[\left\{\epsilon \in [0, \frac{-2(\bar{\gamma}_N - \bar{\gamma}_i)}{\Delta_\gamma^{B,A}}]\right\} \cap \left\{\psi \in [\bar{\gamma}_N - 2(\bar{\gamma}_N - \bar{\gamma}_i), 1]\right\} \cap ...\right.$$

$$\left. ... \cap \left\{c \in \left[0, \frac{nd\frac{(1-\beta)}{\beta}\left[\Delta_\gamma^{B,A}\epsilon + 2(\gamma_N - \gamma_i)\right]}{\bar{\gamma}_N - 2(\bar{\gamma}_N - \bar{\gamma}_i) - \psi}\right]\right\}\right] =$$

$$= \int_0^{\frac{-2(\bar{\gamma}_N - \bar{\gamma}_i)}{\Delta_\gamma^{B,A}}} \int_{\bar{\gamma}_N - 2(\bar{\gamma}_N - \bar{\gamma}_i)}^{1} \int_0^{\frac{nd\frac{(1-\beta)}{\beta}\left[\Delta_\gamma^{B,A}\epsilon + 2(\gamma_N - \gamma_i)\right]}{\bar{\gamma}_N - 2(\bar{\gamma}_N - \bar{\gamma}_i) - \psi}} \frac{1}{c^2} e^{-\frac{1}{c}} \, dc \, d\psi \, d\epsilon$$

Given this scenario, lobby revenue variation located in the specified region will improve individual utility function.

IV) When there is a beneficial lobby maximizer revenue policy ($\psi \in [\bar{\gamma}_N, \bar{\gamma}_N - 2(\bar{\gamma}_N - \bar{\gamma}_i)]$), but a disadvantageous state-probability distortion to the left side ($\epsilon \in [\epsilon_{MIN}, 0]$). This situation will improve individual utility function if $c \in \left[\frac{nd\frac{(1-\beta)}{\beta}\epsilon\Delta_\gamma^{B,A}}{(\bar{\gamma}_N - \psi)}, \infty\right]$. Then, probability of this scenario is:

$$\Pr\left(\{\epsilon \in [\epsilon_{MIN}, 0]\} \cap \{\psi \in [\bar{\gamma}_N, \bar{\gamma}_N - 2(\bar{\gamma}_N - \bar{\gamma}_i)]\} \cap \left\{c \in \left[\frac{nd\frac{(1-\beta)}{\beta}\epsilon\Delta_\gamma^{B,A}}{(\bar{\gamma}_N - \psi)}, \infty\right]\right\}\right)$$

$$= \int_{\epsilon_{MIN}}^{0} \int_{\bar{\gamma}_N}^{\bar{\gamma}_N - 2(\bar{\gamma}_N - \bar{\gamma}_i)} \int_{\frac{nd\frac{(1-\beta)}{\beta}\epsilon\Delta_\gamma^{B,A}}{(\bar{\gamma}_N - \psi)}}^{\infty} \frac{1}{c^2} e^{-\frac{1}{c}} \, dc \, d\psi \, d\epsilon$$

Given this scenario, lobby revenue variation located in the specified region will improve individual utility function.

V) When there is a beneficial lobby maximizer revenue policy ($\psi \in [\bar{\gamma}_N, \bar{\gamma}_N - 2(\bar{\gamma}_N - \bar{\gamma}_i)]$), but a disadvantageous state-probability distortion to the right side

($\epsilon \in [\frac{-2(\bar{\gamma}_N - \bar{\gamma}_i)}{\Delta_\gamma^{B,A}}, \epsilon_{MAX}]$). This situation will improve individual utility function if $c \in \left[\frac{nd\frac{(1-\beta)}{\beta}\left[\Delta_\gamma^{B,A}\epsilon + 2(\gamma_N - \gamma_i)\right]}{\bar{\gamma}_N - 2(\bar{\gamma}_N - \bar{\gamma}_i) - \psi}, \infty\right]$. Then, probability of this scenario is:

$$\Pr\left[\left\{\epsilon \in [\frac{-2(\bar{\gamma}_N - \bar{\gamma}_i)}{\Delta_\gamma^{B,A}}, \epsilon_{MAX}]\right\} \cap \left\{\psi \in [\bar{\gamma}_N, \bar{\gamma}_N - 2(\bar{\gamma}_N - \bar{\gamma}_i)]\right\} \cap ...\right.$$

$$\left. ... \cap \left\{c \in \left[\frac{nd\frac{(1-\beta)}{\beta}\left[\Delta_\gamma^{B,A}\epsilon + 2(\gamma_N - \gamma_i)\right]}{\bar{\gamma}_N - 2(\bar{\gamma}_N - \bar{\gamma}_i) - \psi}, \infty\right]\right\}\right] =$$

$$= \int_{\frac{-2(\bar{\gamma}_N - \bar{\gamma}_i)}{\Delta_\gamma^{B,A}}}^{\epsilon_{MAX}} \int_{\bar{\gamma}_N}^{\bar{\gamma}_N - 2(\bar{\gamma}_N - \bar{\gamma}_i)} \int_{\frac{nd\frac{(1-\beta)}{\beta}\left[\Delta_\gamma^{B,A}\epsilon + 2(\gamma_N - \gamma_i)\right]}{\bar{\gamma}_N - 2(\bar{\gamma}_N - \bar{\gamma}_i) - \psi}}^{\infty} \frac{1}{c^2} e^{-\frac{1}{c}} \, dc \, d\psi \, d\epsilon$$

Given this scenario, lobby revenue variation located in the specified region will improve individual utility function.

VI) When both state-probability distortion (to the left side) and lobby maximizer revenue policy (to the right side) are disadvantageous ($\epsilon \in [\epsilon_{MIN}, 0]$ and $\psi \in [\bar{\gamma}_N - 2(\bar{\gamma}_N - \bar{\gamma}_i), 1]$, respectively), but a certain level of lobby revenue variation weakly increases the individual utility function, $c \in \left[\frac{nd\frac{(1-\beta)}{\beta}\epsilon\Delta_\gamma^{B,A}}{(\bar{\gamma}_N - \psi)}, \frac{nd\frac{(1-\beta)}{\beta}\left[\Delta_\gamma^{B,A}\epsilon + 2(\gamma_N - \gamma_i)\right]}{\bar{\gamma}_N - 2(\bar{\gamma}_N - \bar{\gamma}_i) - \psi}\right]$. Then, probability of this scenario is:

$$\Pr\left[\left\{\epsilon \in [\epsilon_{MIN}, 0]\right\} \cap \left\{\psi \in [\bar{\gamma}_N - 2(\bar{\gamma}_N - \bar{\gamma}_i), 1]\right\} \cap ...\right.$$

$$\left. ... \cap \left\{c \in \left[\frac{nd\frac{(1-\beta)}{\beta}\epsilon\Delta_\gamma^{B,A}}{(\bar{\gamma}_N - \psi)}, \frac{nd\frac{(1-\beta)}{\beta}\left[\Delta_\gamma^{B,A}\epsilon + 2(\gamma_N - \gamma_i)\right]}{\bar{\gamma}_N - 2(\bar{\gamma}_N - \bar{\gamma}_i) - \psi}\right]\right\}\right] =$$

$$= \int_{\epsilon_{MIN}}^{0} \int_{\bar{\gamma}_N - 2(\bar{\gamma}_N - \bar{\gamma}_i)}^{1} \int_{\frac{nd\frac{(1-\beta)}{\beta}\epsilon\Delta_\gamma^{B,A}}{(\bar{\gamma}_N - \psi)}}^{\frac{nd\frac{(1-\beta)}{\beta}\left[\Delta_\gamma^{B,A}\epsilon + 2(\gamma_N - \gamma_i)\right]}{\bar{\gamma}_N - 2(\bar{\gamma}_N - \bar{\gamma}_i) - \psi}} \frac{1}{c^2} e^{-\frac{1}{c}} \, dc \, d\psi \, d\epsilon$$

Given this scenario, lobby revenue variation located in the specified region will improve individual utility function.

VII) When both state-probability distortion (to the right side) and lobby maximizer revenue policy (to the left side) are disadvantageous ($\epsilon \in [\frac{-2(\bar{\gamma}_N - \bar{\gamma}_i)}{\Delta_\gamma^{B,A}}, \epsilon_{MAX}]$ and $\psi \in [0, \bar{\gamma}_N]$, respectively), but a certain level of lobby revenue variation weakly increases the individual utility function, $c \in \left[\frac{nd\frac{(1-\beta)}{\beta}\left[\Delta_\gamma^{B,A}\epsilon + 2(\gamma_N - \gamma_i)\right]}{\bar{\gamma}_N - 2(\bar{\gamma}_N - \bar{\gamma}_i) - \psi}, \frac{nd\frac{(1-\beta)}{\beta}\epsilon\Delta_\gamma^{B,A}}{(\bar{\gamma}_N - \psi)}\right]$. Then, probability of this scenario is:

$$\Pr\left[\left\{\epsilon \in [\frac{-2(\bar{\gamma}_N - \bar{\gamma}_i)}{\Delta_\gamma^{B,A}}, \epsilon_{MAX}]\right\} \cap \left\{\psi \in [0, \bar{\gamma}_N]\right\} \cap ...\right.$$

$$\ldots \cap \left\{ c \in \left[\frac{nd\frac{(1-\beta)}{\beta}\left[\Delta_{\gamma}^{B,A}\epsilon + 2(\gamma_N - \gamma_i)\right]}{\bar{\gamma}_N - 2(\bar{\gamma}_N - \bar{\gamma}_i) - \psi}, \frac{nd\frac{(1-\beta)}{\beta}\epsilon\Delta_{\gamma}^{B,A}}{(\bar{\gamma}_N - \psi)} \right] \right\} \Bigg] =$$

$$= \int_{\frac{-2(\bar{\gamma}_N - \bar{\gamma}_i)}{\Delta_{\gamma}^{B,A}}}^{\epsilon_{MAX}} \int_{0}^{\bar{\gamma}_N} \int_{\frac{nd\frac{(1-\beta)}{\beta}\left[\Delta_{\gamma}^{B,A}\epsilon + 2(\gamma_N - \gamma_i)\right]}{\bar{\gamma}_N - 2(\bar{\gamma}_N - \bar{\gamma}_i) - \psi}}^{\frac{nd\frac{(1-\beta)}{\beta}\epsilon\Delta_{\gamma}^{B,A}}{(\bar{\gamma}_N - \psi)}} \frac{1}{c^2} e^{-\frac{1}{c}} \; dc \; d\psi \; d\epsilon$$

Given this scenario, lobby revenue variation located in the specified region will improve individual utility function.

Then, for those individuals that are characterized for this most preferred policy, their individual trust function will be computed as:

$$\mathbb{T}_N(\bar{\gamma}_i) = \sum_{j=I}^{VII} Pr(\epsilon, \psi, c)$$

d) When $\epsilon_{MAX} \leq \frac{-2(\bar{\gamma}_N - \bar{\gamma}_i)}{\Delta_{\gamma}^{B,A}}$ and:

d.i) $\bar{\gamma}_N \leq \bar{\gamma}_N - 2(\bar{\gamma}_N - \bar{\gamma}_i) < 1$, individual will compute their individual trust through the computation of following probabilities:

I) When state-probability distortion and lobby maximizer revenue policy weakly increases the individual utility function that it will occur if $\epsilon \in [0, \epsilon_{MAX}$ and $\psi \in [\bar{\gamma}_N, \bar{\gamma}_N - 2(\bar{\gamma}_N - \bar{\gamma}_i)]$, respectively. Then, probability of this scenario is:

$$\Pr\left(\{\epsilon \in [0, \epsilon_{MAX}]\} \cap \{\bar{\gamma}_N, \psi \in [\bar{\gamma}_N - 2(\bar{\gamma}_N - \bar{\gamma}_i)]\}\right)$$

$$= \int_{0}^{\epsilon_{MAX}} \int_{\bar{\gamma}_N}^{\bar{\gamma}_N - 2(\bar{\gamma}_N - \bar{\gamma}_i)} 1 \; d\psi \; d\epsilon$$

Notice that this situation does not require of cost bounds since both parameters will increase individual utility function regardless the cost level.

II) When there is a beneficial state-probability distortion ($\epsilon \in [0, \epsilon_{MAX}]$), but a disadvantageous lobby maximizer revenue policy to the left side ($\psi \in [0, \bar{\gamma}_N]$). This situation will improve individual utility function if $c \in \left[0, \frac{nd\frac{(1-\beta)}{\beta}\epsilon\Delta_{\gamma}^{B,A}}{(\bar{\gamma}_N - \psi)}\right]$. Then, probability of this scenario is:

$$\Pr\left(\{\epsilon \in [0, \epsilon_{MAX}]\} \cap \{\psi \in [0, \bar{\gamma}_N]\} \cap \left\{ c \in \left[0, \frac{nd\frac{(1-\beta)}{\beta}\epsilon\Delta_{\gamma}^{B,A}}{(\bar{\gamma}_N - \psi)}\right] \right\} \right)$$

$$= \int_{0}^{\epsilon_{MAX}} \int_{0}^{\bar{\gamma}_N} \int_{0}^{\frac{nd^{\frac{(1-\beta)}{\beta}} \epsilon \Delta_{\gamma}^{B,A}}{(\bar{\gamma}_N - \psi)}} \frac{1}{c^2} e^{-\frac{1}{c}} \ dc \ d\psi \ d\epsilon$$

Given this scenario, lobby revenue variation located in the specified region will improve individual utility function.

III) When there is a beneficial state-probability distortion ($\epsilon \in [0, \epsilon_{MAX}]$), but a disadvantageous lobby maximizer revenue policy to the right side ($\psi \in [\bar{\gamma}_N - 2(\bar{\gamma}_N - \bar{\gamma}_i), 1]$). This situation will improve individual utility function if $c \in \left[0, \frac{nd^{\frac{(1-\beta)}{\beta}} \left[\Delta_{\gamma}^{B,A} \epsilon + 2(\gamma_N - \gamma_i)\right]}{\bar{\gamma}_N - 2(\bar{\gamma}_N - \bar{\gamma}_i) - \psi}\right]$. Then, probability of this scenario is:

$$\Pr\left[\left\{\epsilon \in [0, \epsilon_{MAX}]\right\} \cap \left\{\psi \in [\bar{\gamma}_N - 2(\bar{\gamma}_N - \bar{\gamma}_i), 1]\right\} \cap ...\right.$$

$$\left. ... \cap \left\{c \in \left[0, \frac{nd^{\frac{(1-\beta)}{\beta}} \left[\Delta_{\gamma}^{B,A} \epsilon + 2(\gamma_N - \gamma_i)\right]}{\bar{\gamma}_N - 2(\bar{\gamma}_N - \bar{\gamma}_i) - \psi}\right]\right\}\right] =$$

$$= \int_{0}^{\epsilon_{MAX}} \int_{\bar{\gamma}_N - 2(\bar{\gamma}_N - \bar{\gamma}_i)}^{1} \int_{0}^{\frac{nd^{\frac{(1-\beta)}{\beta}} \left[\Delta_{\gamma}^{B,A} \epsilon + 2(\gamma_N - \gamma_i)\right]}{\bar{\gamma}_N - 2(\bar{\gamma}_N - \bar{\gamma}_i) - \psi}} \frac{1}{c^2} e^{-\frac{1}{c}} \ dc \ d\psi \ d\epsilon$$

Given this scenario, lobby revenue variation located in the specified region will improve individual utility function.

IV) When there is a beneficial lobby maximizer revenue policy ($\psi \in [\bar{\gamma}_N, \bar{\gamma}_N - 2(\bar{\gamma}_N - \bar{\gamma}_i)]$), but a disadvantageous state-probability distortion to the left side ($\epsilon \in [\epsilon_{MIN}, 0]$). This situation will improve individual utility function if $c \in \left[\frac{nd^{\frac{(1-\beta)}{\beta}} \epsilon \Delta_{\gamma}^{B,A}}{(\bar{\gamma}_N - \psi)}, \infty\right]$. Then, probability of this scenario is:

$$\Pr\left(\{\epsilon \in [\epsilon_{MIN}, 0]\} \cap \{\psi \in [\bar{\gamma}_N, \bar{\gamma}_N - 2(\bar{\gamma}_N - \bar{\gamma}_i)]\} \cap \left\{c \in \left[\frac{nd^{\frac{(1-\beta)}{\beta}} \epsilon \Delta_{\gamma}^{B,A}}{(\bar{\gamma}_N - \psi)}, \infty\right]\right\}\right)$$

$$= \int_{\epsilon_{MIN}}^{0} \int_{\bar{\gamma}_N}^{\bar{\gamma}_N - 2(\bar{\gamma}_N - \bar{\gamma}_i)} \int_{\frac{nd^{\frac{(1-\beta)}{\beta}} \epsilon \Delta_{\gamma}^{B,A}}{(\bar{\gamma}_N - \psi)}}^{\infty} \frac{1}{c^2} e^{-\frac{1}{c}} \ dc \ d\psi \ d\epsilon$$

Given this scenario, lobby revenue variation located in the specified region will improve individual utility function.

V) When both state-probability distortion (to the left side) and lobby maximizer revenue policy (to the right side) are disadvantageous ($\epsilon \in [\epsilon_{MIN}, 0]$ and $\psi \in [\bar{\gamma}_N - 2(\bar{\gamma}_N - \bar{\gamma}_i), 1]$, respectively), but a certain level of lobby revenue variation weakly increases the individual utility function, $c \in \left[\frac{nd^{\frac{(1-\beta)}{\beta}} \epsilon \Delta_{\gamma}^{B,A}}{(\bar{\gamma}_N - \psi)}, \frac{nd^{\frac{(1-\beta)}{\beta}} \left[\Delta_{\gamma}^{B,A} \epsilon + 2(\gamma_N - \gamma_i)\right]}{\bar{\gamma}_N - 2(\bar{\gamma}_N - \bar{\gamma}_i) - \psi}\right]$. Then, probability of this scenario is:

$$\Pr\left[\Big\{\epsilon \in [\epsilon_{MIN}, 0]\Big\} \cap \Big\{\psi \in [\bar{\gamma}_N - 2(\bar{\gamma}_N - \bar{\gamma}_i), 1]\Big\} \cap ...\right.$$

$$\left. ... \cap \left\{ c \in \left[\frac{nd\frac{(1-\beta)}{\beta}\epsilon\Delta_{\gamma}^{B,A}}{(\bar{\gamma}_N - \psi)}, \frac{nd\frac{(1-\beta)}{\beta}\left[\Delta_{\gamma}^{B,A}\epsilon + 2(\gamma_N - \gamma_i)\right]}{\bar{\gamma}_N - 2(\bar{\gamma}_N - \bar{\gamma}_i) - \psi} \right] \right\} \right] =$$

$$= \int_{\epsilon_{MIN}}^{0} \int_{\bar{\gamma}_N - 2(\bar{\gamma}_N - \bar{\gamma}_i)}^{1} \int_{\frac{nd\frac{(1-\beta)}{\beta}\epsilon\Delta_{\gamma}^{B,A}}{(\bar{\gamma}_N - \psi)}}^{\frac{nd\frac{(1-\beta)}{\beta}\left[\Delta_{\gamma}^{B,A}\epsilon + 2(\gamma_N - \gamma_i)\right]}{\bar{\gamma}_N - 2(\bar{\gamma}_N - \bar{\gamma}_i) - \psi}} \frac{1}{c^2} e^{-\frac{1}{c}} \; dc \; d\psi \; d\epsilon$$

Given this scenario, lobby revenue variation located in the specified region will improve individual utility function.

Then, for those individuals that are characterized for this most preferred policy, their individual trust function will be computed as:

$$\mathbb{T}_N(\bar{\gamma}_i) = \sum_{j=I}^{V} Pr(\epsilon, \psi, c)$$

d.ii) $\bar{\gamma}_N - 2(\bar{\gamma}_N - \bar{\gamma}_i) \geq 1$, individual will compute their individual trust through the computation of following probabilities:

I) When state-probability distortion and lobby maximizer revenue policy weakly increases the individual utility function that it will occur if $\epsilon \in [0, \epsilon_{MAX}$ and $\psi \in [\bar{\gamma}_N, 1]$, respectively. Then, probability of this scenario is:

$$\Pr\left(\{\epsilon \in [0, \epsilon_{MAX}]\} \cap \{\bar{\gamma}_N, 1\}\right) = \int_{0}^{\epsilon_{MAX}} \int_{\bar{\gamma}_N}^{\bar{\gamma}_N - 2(\bar{\gamma}_N - \bar{\gamma}_i)} 1 \; d\psi \; d\epsilon$$

Notice that this situation does not require of cost bounds since both parameters will increase individual utility function regardless the cost level.

II) When there is a beneficial state-probability distortion ($\epsilon \in [0, \epsilon_{MAX}]$), but a disadvantageous lobby maximizer revenue policy to the left side ($\psi \in [0, \bar{\gamma}_N]$). This situation will improve individual utility function if $c \in \left[0, \frac{nd\frac{(1-\beta)}{\beta}\epsilon\Delta_{\gamma}^{B,A}}{(\bar{\gamma}_N - \psi)}\right]$. Then, probability of this scenario is:

$$\Pr\left(\{\epsilon \in [0, \epsilon_{MAX}]\} \cap \{\psi \in [0, \bar{\gamma}_N]\} \cap \left\{ c \in \left[0, \frac{nd\frac{(1-\beta)}{\beta}\epsilon\Delta_{\gamma}^{B,A}}{(\bar{\gamma}_N - \psi)}\right] \right\}\right)$$

$$= \int_{0}^{\epsilon_{MAX}} \int_{0}^{\bar{\gamma}_N} \int_{0}^{\frac{nd\frac{(1-\beta)}{\beta}\epsilon\Delta_{\gamma}^{B,A}}{(\bar{\gamma}_N-\psi)}} \frac{1}{c^2}e^{-\frac{1}{c}} \; dc \; d\psi \; d\epsilon$$

Given this scenario, lobby revenue variation located in the specified region will improve individual utility function.

III) When there is a beneficial lobby maximizer revenue policy ($\psi \in [\bar{\gamma}_N, 1]$), but a disadvantageous state-probability distortion to the left side ($\epsilon \in [\epsilon_{MIN}, 0]$). This situation will improve individual utility function if $c \in \left[\frac{nd\frac{(1-\beta)}{\beta}\epsilon\Delta_{\gamma}^{B,A}}{(\bar{\gamma}_N-\psi)}, \infty\right]$. Then, probability of this scenario is:

$$\Pr\left(\{\epsilon \in [\epsilon_{MIN}, 0]\} \cap \{\psi \in [\bar{\gamma}_N, 1]\} \cap \left\{c \in \left[\frac{nd\frac{(1-\beta)}{\beta}\epsilon\Delta_{\gamma}^{B,A}}{(\bar{\gamma}_N - \psi)}, \infty\right]\right\}\right)$$

$$= \int_{\epsilon_{MIN}}^{0} \int_{\bar{\gamma}_N}^{1} \int_{\frac{nd\frac{(1-\beta)}{\beta}\epsilon\Delta_{\gamma}^{B,A}}{(\bar{\gamma}_N-\psi)}}^{\infty} \frac{1}{c^2}e^{-\frac{1}{c}} \; dc \; d\psi \; d\epsilon$$

Given this scenario, lobby revenue variation located in the specified region will improve individual utility function.

Then, for those individuals that are characterized for this most preferred policy, their individual trust function will be computed as:

$$\mathbb{T}_N(\bar{\gamma}_i) = \sum_{j=I}^{III} Pr(\epsilon, \psi, c)$$

Appendix: Chapter 2

Competitive Equilibrium in Perfect Competitive Markets

Consider the maximization problem mentioned in Equation 2.2 and take the first order condition with respect to x_t and y_t, since a weakly positive z_t will absorbe the income effect generated in the optimisation problem. Then, for all $t \in \{1, 2\}$:

$$(1-\theta)\left[a - bx_t\right] - y_t - p_{x,t} \leq 0,\ x_t \geq 0,\ \text{c.s.}$$

$$\theta\left[a - by_t\right] - x_t - p_{y,t} \leq 0,\ y_t \geq 0,\ \text{c.s.}$$

Resulting in the following inverse demand functions:

$$p_{x,t} = (1-\theta)[a - bx_t] - y_t$$

$$p_{y,t} = \theta[a - by_t] - x_t$$

Now, consider the optimisation problem of multi-supplier firms that perfectly competes, being price-takers, in the market, expressed in Equation 2.1. By analysing the first order conditions:

$$p_{x,t} - c_0 + c_1 r_t - \lambda \leq 0,\ x_t \geq 0,\ \text{c.s.}$$

$$p_{y,t} - k_0 - k_1 y_t \leq 0,\ y_t \geq 0,\ \text{c.s.}$$

$$r_0 \geq x_t + x_{t+1},\ \lambda \geq 0,\ \text{c.s.}$$

We observe that the type of energy supplied in the market will depend on some parameters:

- **No production of energy** ($x_t = 0$ and $y_t = 0$): Consequently, $\lambda = 0$. It will occurs if $(1-\theta)a < c_0 - c_1 r_0$ and $\theta a < k_0$. Specifically, if $\theta \in \left(\frac{a - c_0 + c_1 r_0}{a}, \frac{k_0}{a}\right)$.

- **Only production of green energy** ($x_t = 0$ and $y_t > 0$): Consequently, $\lambda = 0$. By respecting the FOC of green energy, we identify that $y_t = \frac{\theta a - k_0}{\theta b + k_1}$. It will occurs if $(1-\theta)a - y_t < c_0 - c_1 r_0$. Specifically, previous condition could be understood as the minimum level of green energy such that brown energy is not economically beneficial, $y_t > (1-\theta)a - c_0 + c_1 r_0$.

- **Production of brown energy**: However, the FOC of the shadow price, λ, could be respected or not, suppose the following cases:

- **Only partial exhaustion of brown energy** ($x_t > 0$, $\lambda = 0$ and $y_t = 0$): Notice $r_0 \geq x_t + x_{t+1}$. By respecting the FOC of brown energy, we identify that for current (t) and future ($t+1$) periods:

$$x_t = \frac{(1-\theta)a - [c_0 - c_1 r_0]}{(1-\theta)b + c_1}$$

$$x_{t+1} = \frac{(1-\theta)a - \left[c_0 - c_1 \left(r_0 - \frac{(1-\theta)a - [c_0 - c_1 r_0]}{(1-\theta)b + c_1}\right)\right]}{(1-\theta)b + c_1}$$

where $r_t = r_0 - x_t$ and $r_{t+1} = r_0 - x_t - x_{t+1}$. Notice that future periods could be expressed as:

$$x_{t+1} = \frac{(1-\theta)b}{(1-\theta)b + c_1} x_t$$

Notice that it will occur if $\theta a - x_t < k_0$. Specifically, previous condition could be understood as the minimum level of brown energy such that green energy is not economically beneficial, $x_t > \theta a - k_0$.

- **Only full exhaustion of brown energy** ($x_t > 0$, $\lambda > 0$ and $y_t = 0$): Notice $r_0 = x_t + x_{t+1}$. By respecting the FOC of brown energy and shadow price, we identify that for current (t) and future ($t+1$) periods needs to respect the following Hotelling rule:

$$p_{x,t} - c_0 + c_1 r_t = p_{x,t+1} - c_0 + c_1 r_{t+1}$$

Since $r_0 = x_t + x_{t+1}$:

$$x_t = \frac{(1-\theta)b + c_1}{2(1-\theta)b + c_1} r_0$$

$$x_{t+1} = \frac{(1-\theta)b}{2(1-\theta)b + c_1} r_0$$

Notice that it will also occur if current level of brown energy exceeds the minimum one such that green energy is not economically beneficial, $x_t > \theta a - k_0$.

- **Simultaneous supply of both types of energy with partial exhaustion of brown energy** ($x_t > 0$, $\lambda = 0$ and $y_t > 0$): Notice $r_0 > x_t + x_{t+1}$. The FOC of green energy allow us to understand that green energy will behave for all periods in the following way:

$$y_t = \frac{\theta a - x_t - k_0}{\theta b + k_1}$$

Whereas the brown energy will respect the FOC, where the fossil fuel supply for current (t) and future ($t+1$) periods will be equivalent to:

$$x_t = \frac{[\theta b + k_1]\,[(1-\theta)a - c_0 + c_1 r_0] - [\theta a - k_0]}{[\theta b + k_1]\,[(1-\theta)b + c_1] - 1}$$

$$x_{t+1} = \frac{[\theta b + k_1]\,[(1-\theta)b] - 1}{[\theta b + k_1]\,[(1-\theta)b + c_1] - 1} x_t$$

- **Simultaneous supply of both types of energy with full exhaustion of brown energy** ($x_t > 0$, $\lambda > 0$ and $y_t > 0$): Notice $r_0 = x_t + x_{t+1}$. The FOC of green energy allow us to understand that green energy will behave for all periods in the following way:

  $$y_t = \frac{\theta a - x_t - k_0}{\theta b + k_1}$$

  Whereas the brown energy will respect both shadow price and their own FOC by satisfying the Hotelling rule:

  $$p_{x,t} - c_0 + c_1 r_t = p_{x,t+1} - c_0 + c_1 r_{t+1}$$

  Then, the fossil fuel supply for current (t) and future ($t+1$) periods will be equivalent to:

  $$x_t = \frac{[(1-\theta)b(\theta b + k_1) - 1] + c_1(\theta b + k_1)}{2\,[(1-\theta)b(\theta b + k_1) - 1] + c_1(\theta b + k_1)} r_0$$

  $$x_{t+1} = \frac{[(1-\theta)b(\theta b + k_1) - 1]}{2\,[(1-\theta)b(\theta b + k_1) - 1] + c_1(\theta b + k_1)} r_0$$

Competitive Equilibrium in Monopoly

Consider the maximization problem mentioned described in previous proof; analyse the optimisation problem of a multi-supplier firm who competes in a monopoly, being price-makers, in the market, expressed in Equation 2.1. By analysing the first order conditions:

$$p_{x,t} - (1-\theta)bx_t - y_t - c_0 + c_1 r_t \leq 0,\ x_t - \lambda \geq 0,\ \text{c.s.}$$

$$p_{y,t} - \theta b y_t - x_t - k_0 - k_1 y_t \leq 0,\ y_t \geq 0,\ \text{c.s.}$$

$$r_0 \geq x_t + x_{t+1},\ \lambda \geq 0,\ \text{c.s.}$$

Notice that the marginal revenue of brown energy is equivalent to $MR_x = p_{x,t} - (1 - \theta)bx_t - y_t = 2p_{x,t} - (1-\theta)a$ and $MR_y = p_{y,t} - \theta b y_t - x_t = 2p_{y,t} - \theta a$ We observe that the type of energy supplied in the market will depend on some parameters:

- **No production of energy** ($x_t = 0$ and $y_t = 0$): Consequently, $\lambda = 0$. It will occurs if $(1-\theta)a < c_0 - c_1 r_0$ and $\theta a < k_0$. Specifically, if $\theta \in \left(\frac{a-c_0+c_1r_0}{a}, \frac{k_0}{a}\right)$.
- **Only production of green energy** ($x_t = 0$ and $y_t > 0$): Consequently, $\lambda = 0$. By respecting the FOC of green energy, we identify that $y_t = \frac{\theta a - k_0}{2\theta b + k_1}$. It will occurs if $(1-\theta)a - 2y_t < c_0 - c_1 r_0$. Specifically, previous condition could be understood as the minimum level of green energy such that brown energy is not economically beneficial, $y_t > \frac{(1-\theta)a - c_0 + c_1 r_0}{2}$.
- **Production of brown energy**: However, the FOC of the shadow price, λ, could be respected or not, suppose the following cases:
  - **Only partial exhaustion of brown energy** ($x_t > 0$, $\lambda = 0$ and $y_t = 0$): Notice $r_0 > x_t + x_{t+1}$. By respecting the FOC of brown energy, we identify that for current (t) and future ($t+1$) periods:

    $$x_t = \frac{(1-\theta)a - c_0 + c_1 r_0}{2(1-\theta)b + c_1}$$

    $$x_{t+1} = \frac{(1-\theta)a - c_0 + c_1(r_0 - x_t)}{2(1-\theta)b + c_1}$$

    where $r_t = r_0 - x_t$ and $r_{t+1} = r_0 - x_t - x_{t+1}$. Notice that future periods could be expressed as:

    $$x_{t+1} = \frac{2(1-\theta)b}{2(1-\theta)b + c_1} x_t$$

    Notice that it will occur if $\theta a - 2x_t < k_0$. Specifically, previous condition could be understood as the minimum level of brown energy such that green energy is not economically beneficial, $x_t > \frac{\theta a - k_0}{2}$.
  - **Only full exhaustion of brown energy** ($x_t > 0$, $\lambda > 0$ and $y_t = 0$): Notice $r_0 = x_t + x_{t+1}$. By respecting the FOC of brown energy and shadow price, we identify that for current (t) and future ($t+1$) periods needs to respect the following hotelling rule:

    $$p_{x,t} - (1-\theta)bx_t - c_0 + c_1 r_t = p_{x,t+1} - (1-\theta)bx_{t+1} - c_0 + c_1 r_{t+1}$$

Since $r_0 = x_t + x_{t+1}$:

$$x_t = \frac{2(1-\theta)b + c_1}{4(1-\theta)b + c_1} r_0$$

$$x_{t+1} = \frac{2(1-\theta)b}{4(1-\theta)b + c_1} r_0$$

Notice that it will also occur if current level of brown energy exceeds the minimum one such that green energy is not economically beneficial, $x_t > \frac{\theta a - k_0}{2}$.

- **Simultaneous supply of both types of energy with partial exhaustion of brown energy** ($x_t > 0$, $\lambda = 0$ and $y_t > 0$): Notice $r_0 > x_t + x_{t+1}$. The FOC of green energy allow us to understand that green energy will behave for all periods in the following way:

  $$y_t = \frac{\theta a - 2x_t - k_0}{2\theta b + k_1}$$

  Whereas the brown energy will respect the FOC, where the fossil fuel supply for current (t) and future ($t+1$) periods will be equivalent to:

  $$x_t = \frac{[2\theta b + k_1]\,[(1-\theta)a - c_0 + c_1 r_0] - 2[\theta a - k_0]}{\Big[[2\theta b + k_1]\Big(2(1-\theta)b + c_1\Big) - 4\Big]}$$

  $$x_{t+1} = \left(\frac{\Big[[2\theta b + k_1]\Big(2(1-\theta)b\Big) - 4\Big]}{\Big[[2\theta b + k_1]\Big(2(1-\theta)b + c_1\Big) - 4\Big]}\right) x_t$$

- **Simultaneous supply of both types of energy with full exhaustion of brown energy** ($x_t > 0$, $\lambda > 0$ and $y_t > 0$): Notice $r_0 = x_t + x_{t+1}$. The FOC of green energy allow us to understand that green energy will behave for all periods in the following way:

  $$y_t = \frac{\theta a - 2x_t - k_0}{2\theta b + k_1}$$

  Whereas the brown energy will respect both shadow price and their own FOC by satisfying the Hotelling rule:

  $$p_{x,t} - (1-\theta)bx_t - y_t - c_0 + c_1 r_t = p_{x,t+1} - (1-\theta)bx_{t+1} - y_{t+1} - c_0 + c_1 r_{t+1}$$

  Then, the fossil fuel supply for current (t) and future ($t+1$) periods will be equivalent to:

$$x_t = \frac{2\Big[(1-\theta)b[2\theta b + k_1] - 2\Big] + c_1[2\theta b + k_1]}{4\Big[(1-\theta)b[2\theta b + k_1] - 2\Big] + c_1[2\theta b + k_1]} r_0$$

$$x_{t+1} = \frac{2\Big[(1-\theta)b[2\theta b + k_1] - 2\Big]}{4\Big[(1-\theta)b[2\theta b + k_1] - 2\Big] + c_1[2\theta b + k_1]} r_0$$

Proof

Proof 4 *Consider all the market types, perfect competition and monopoly, and suppose that all of them do not bind the fossil fuel resource constraint. Now, let's start considering that the preference parameter is lower enough, call it* $\underline{\theta}$*, such that only brown energy is produced, i.e. Equation 2.7 is satisfied for perfect competitive markets, whereas Equation 2.11 is satisfied for monopoly. Then, any marginal increase in the preference parameter will reduce the amount of fossil fuel in perfect competitive markets:*

$$\frac{\partial x_1^{PC}(\underline{\theta})}{\partial \theta} < 0 \qquad \frac{\partial x_2^{PC}(\underline{\theta})}{\partial \theta} < 0$$

With the same effects in monopoly:

$$\frac{\partial x_1^{M}(\underline{\theta})}{\partial \theta} < 0 \qquad \frac{\partial x_2^{M}(\underline{\theta})}{\partial \theta} < 0$$

Now, consider that both types of markets will face a higher enough preference parameter, call it $\bar{\theta}$*, such that both brown and green energy are simultaneously produced, i.e. Equation 2.7 is not satisfied for perfect competitive markets, whereas Equation 2.11 is not satisfied for monopoly. Then, any marginal increase in the preference parameter will also reduce the amount of fossil fuel in perfect competitive markets:*

$$\frac{\partial x_1^{PC}(\bar{\theta})}{\partial \theta} < 0 \qquad \frac{\partial x_2^{PC}(\bar{\theta})}{\partial \theta} < 0$$

With the same effects in monopoly:

$$\frac{\partial x_1^{M}(\bar{\theta})}{\partial \theta} < 0 \qquad \frac{\partial x_2^{M}(\bar{\theta})}{\partial \theta} < 0$$

Additionally, under the presence of green energy, the amount of brown energy in current periods will increase in equilibrium in both types of markets:

$$\frac{\partial x_1^{PC}(\bar{\theta})}{\partial k_1} > 0 \qquad \frac{\partial x_1^{M}(\bar{\theta})}{\partial k_1} > 0$$

However, consider these levels of the preference parameter degree such that reaches the equality in Equation 2.7 in perfect competitive market (named as $\tilde{\theta}^{PC}$) and the equality in Equation 2.11 in monopoly (named as $\tilde{\theta}^{M}$). Around this value, green energy will be introduced into the market. Suppose some variation of the preference parameter that moves from the right side of previous level to the left side. Notice that under $k_1 = 0$, the following is true for both types of markets:

$$\lim_{\theta \to \tilde{\theta}^{PC-}} x_1^{PC}(\theta) = \lim_{\theta \to \tilde{\theta}^{PC+}} x_1^{PC}(\theta)$$

$$\lim_{\theta \to \tilde{\theta}^{M-}} x_1^{M}(\theta) = \lim_{\theta \to \tilde{\theta}^{M+}} x_1^{M}(\theta)$$

But, as mentioned previously, under increases of k_1, the level of brown under the presence of the green energy will increase. Then, under this variation, there will be an increase in current periods. With respect to the brown energy provided in future periods, notice that there is a reduction compared to the current period production. However, under the presence of the green energy, the drop in future periods will be smaller. For example, the drop of the brown energy in perfect competition moves from:

$$\frac{(1-\theta)b}{(1-\theta)b + c_1} \rightarrow \frac{[\theta b + k_1]\,[(1-\theta)b] - 1}{[\theta b + k_1]\,[(1-\theta)b + c_1] - 1}$$

Whereas in monopoly moves from:

$$\frac{2(1-\theta)b}{2(1-\theta)b + c_1} \rightarrow \frac{[2\theta b + k_1](2(1-\theta)b) - 4}{[2\theta b + k_1](2(1-\theta)b + c_1) - 4}$$

Notice that it will be possible to increase the extractions in that period under lower levels of k_1. Then, there will be a rise in the amount of resources extracted at the end of the second period, under sufficiently lower levels of k_1.

Proof 5 *Consider all the market types, perfect competition and monopoly, and suppose that all of them bind the fossil fuel resource constraint. Now, let's start considering that the preference parameter is lower enough, call it $\underline{\theta}$, such that only brown energy is produced, i.e. Equation 2.9 is satisfied for perfect competitive markets, whereas Equation 2.13 is satisfied for monopoly. Then, any marginal increase in the preference parameter will* *reduce the amount of fossil fuel in perfect competitive markets:*

$$\frac{\partial x_1^{PC}(\underline{\theta})}{\partial \theta} > 0 \qquad \frac{\partial x_2^{PC}(\underline{\theta})}{\partial \theta} < 0$$

With the same effects in monopoly:

$$\frac{\partial x_1^{M}(\underline{\theta})}{\partial \theta} > 0 \qquad \frac{\partial x_2^{M}(\underline{\theta})}{\partial \theta} < 0$$

Now, consider that both types of markets will face a higher enough preference parameter, call it $\bar{\theta}$, such that both brown and green energy are simultaneously produced, i.e. Equation 2.9 is not satisfied for perfect competitive markets, whereas Equation 2.13 is not satisfied for monopoly. Then, any marginal increase in the preference parameter will also reduce the amount of fossil fuel in perfect competitive markets:

$$\frac{\partial x_1^{PC}(\bar{\theta})}{\partial \theta} > 0 \qquad \frac{\partial x_2^{PC}(\bar{\theta})}{\partial \theta} < 0$$

With the same effects in monopoly:

$$\frac{\partial x_1^{M}(\bar{\theta})}{\partial \theta} > 0 \qquad \frac{\partial x_2^{M}(\bar{\theta})}{\partial \theta} < 0$$

Now, consider these levels of the preference parameter degree such that reaches the equality in Equation 2.9 in perfect competitive market (named as $\tilde{\theta}^{PC}$) and the equality in Equation 2.13 in monopoly (named as $\tilde{\theta}^{M}$). Around this value, green energy will be introduced into the market. Suppose some variation of the preference parameter that moves from the right side of previous level to the left side. Then:

$$\lim_{\theta \to \tilde{\theta}^{PC-}} x_1^{PC}(\theta) < \lim_{\theta \to \tilde{\theta}^{PC+}} x_1^{PC}(\theta)$$

$$\lim_{\theta \to \tilde{\theta}^{M-}} x_1^{M}(\theta) < \lim_{\theta \to \tilde{\theta}^{M+}} x_1^{M}(\theta)$$

In both markets structure, the introduction of the green energy in the market will motivate an increase in the amount of brown energy in current periods in the competitive equilibrium. Since the available resources are fully exhausted, then there will be also a reduction in the amount of brown energy provided in future periods in the competitive equilibrium.

Appendix: Chapter 3

Proof

Proof 6 *The optimisation problem could be reduced to:*

$$\begin{aligned} \max_{y_i,\theta_i} \quad & B_i\left((\theta_i R)^{\phi}, y_i\right) - C\left((\theta_i R)^{\phi}\right) - K(y_i) - D(E_N) \\ s.t. \quad & y_i \geq 0, \\ & 1 \geq \theta_i \geq 0 \end{aligned}$$

Consider the non-cooperative maximization problem of a country i:

$$\mathcal{L}(y_i, \theta_i, \lambda) = a_i^x (\theta_i R)^{\phi} - \frac{(\theta_i R)^{2\phi}}{2} + a^y y_i - \frac{y_i^2}{2} - (\theta_i R)^{\phi} y_i - \frac{c\,(\theta_i R)^{2\phi}}{2} - k\frac{y_i^2}{2} - d\delta\left(\sum_{i=1}^{N} x_i\right) - \lambda(\theta_i - 1)$$

The Kuhn-Tucker conditions are:

$$\frac{\delta \mathcal{L}}{\delta y_i} : a^y - (1+k)y_i - (\theta_i R)^{\phi} \leq 0 \qquad y_i \geq 0 \tag{7}$$

$$\frac{\delta \mathcal{L}}{\delta \theta_i} : (a_i^x - y_i - d\delta)\,\phi\theta_i^{\phi-1} R^{\phi} - (1+c)\,\phi\theta_i^{2\phi-1} R^{2\phi} - \lambda \leq 0 \qquad \theta_i \geq 0 \tag{8}$$

$$\frac{\delta \mathcal{L}}{\delta \lambda} : \theta_i - 1 \leq 0 \qquad \lambda \geq 0 \tag{9}$$

By analysing the Kuhn-Tucker conditions, we detect:

Case I: Suppose the situation where $\theta_i = 0$ and $y_i = 0$. From Equation 9, $\lambda = 0$; and Equation 7 results in $a^y < 0$, being not possible by construction. Additionally, Equation 8 becomes:

$$\frac{a_i^x - d\delta}{1+c} < (\theta_i R)^{\phi}$$

Being not possible since $a_i^x \geq d\delta$.

Case II: Suppose the situation where $\theta_i = 0$ and $y_i > 0$. From Equation 9, $\lambda = 0$. In Equation 7, $y_i = \frac{a^y}{1+k}$. Last, from Equation 8:

$$\frac{\left(a_i^x - \frac{a^y}{1+k} - d\delta\right)}{(1+c)} < (\theta_i R)^{\phi}$$

Being possible for those countries that satisfies $a_i^x < \frac{a^y}{1+k} + d\delta$ since $W_i^{NC}(\theta_i = 0, y_i > 0) = \frac{1}{2}\frac{(a^y)^2}{1+k} - -d\delta\sum_{-i\in N/\{i\}} x_{-i} > -d\delta\sum_{-i\in N/\{i\}} x_{-i} = W_i^{NC}(\theta_i = 0, y_i = 0)$.

Case III: Suppose the situation where $\theta_i > 0$ and $y_i = 0$. Then, two potential scenarios emerge from Equation 9: $0 < \theta_i < 1$ (and $\lambda = 0$) or $\theta_i = 1$ (and $\lambda > 0$).

- Assume $0 < \theta_i < 1$. From Equation 8, $\theta_i = \frac{\left(a_i^x - d\delta\right)^{\frac{1}{\phi}}}{(1+c)^{\frac{1}{\phi}} R}$. If it is combined in Equation 7, then it will result in $(1+c)a^y + d\delta < a_i^x$, being not possible since $a^y > a_i^x$.
- Assume $\theta_i = 1$. By combining Equation 8 and 9:

  $$\phi R^{\phi} \left\{ \left(a_j^x - d\delta\right) - (1+c)\left(R\right)^{\phi} \right\} > 0$$

  Being possible if $\frac{a_i^x - d\delta}{1+c} > R^{\phi}$. Additionally, Equation 7 results in $a^y < R^{\phi}$. By combining both, it will end up in the following condition, $a_i^x > d\delta + (1+c)a^y$, being not possible by construction, since $a^y > a_i^x$.

Case IV: Suppose the situation where $\theta_i > 0$ and $y_i > 0$. Then, two potential scenarios emerge from Equation 9: $1 > \theta_i > 0$ (and $\lambda = 0$) or $\theta_i = 1$ (and $\lambda > 0$).

- Assume $0 < \theta_i < 1$. From Equation 8, $\theta_i = \frac{\left(a_i^x - y_i - d\delta\right)^{\frac{1}{\phi}}}{(1+c)^{\frac{1}{\phi}} R}$. In addition, Equation 7 shows that $y_i = \frac{a^y - (\theta_i R)^{\phi}}{(1+k)}$ When both are combined, then it will result in:

  $$\theta_i = \frac{\left((1+k)\left(a_i^x - d\delta\right) - a^y\right)^{\frac{1}{\phi}}}{\left((1+k)(1+c) - 1\right)^{\frac{1}{\phi}} R} \qquad y_i = \frac{(1+c)a^y - \left(a_i^x - d\delta\right)}{(1+k)(1+c) - 1}$$

  Resulting in the following welfare function:

  $$W_i^{NC}\left(a_i^x \middle| \theta_i \in (0,1), y_i > 0\right) =$$

  $$= \frac{(1+k)(a_i^x - d\delta)^2 + (1+c)(a^y)^2 - 2a^y(a_i^x - d\delta)}{2\left[(1+k)(1+c) - 1\right]} - d\delta \sum_{-i \in N/\{i\}} x_{-i}$$

- Assume $\theta_i = 1$. From Equation 8, $R \leq \left(\frac{a_i^x - y - d\delta}{(1+c)}\right)^{\frac{1}{\phi}}$. Additionally, Equation 7 results in $y_i = \frac{a^y - R^{\phi}}{(1+k)}$. By combining both previous conditions with Equation 9, $R \leq \left(\frac{(1+k)\left(a_i^x - d\delta\right) - a^y}{((1+k)(1+c) - 1)}\right)^{\frac{1}{\phi}}$. Since R is symmetric among all countries, the level of resources will be equivalent to $R = \left(\frac{(1+k)(\tilde{a}^x - d\delta) - a^y}{((1+k)(1+c) - 1)}\right)^{\frac{1}{\phi}}$, where $\tilde{a}^x$ is the minimum level of the dirty energy preference parameter such that a country will full-exhaust their resources. By using previous resource level in the optimal level of *resources, we identify:*

$$x_i = R^{\phi} = \frac{(1+k)\left(\tilde{a}^x - d\delta\right) - a^y}{((1+k)(1+c)-1)}$$

And:

$$y_i = \frac{a^y - R^{\phi}}{(1+k)} = \frac{(1+c)a^y - \left(\tilde{a}^x - d\delta\right)}{(1+k)(1+c)-1}$$

With a welfare level equivalent to:

$$W_i^{NC}\left(a_i^x \middle| \theta_i = 1, y_i > 0\right) = \frac{(1+k)(a_i^x - d\delta)(\tilde{a}^x - d\delta) + (1+c)(a^y)^2}{2\left[(1+k)(1+c)-1\right]} +$$

$$+ \frac{-(1+k)(\tilde{a}^x - d\delta)\left[\tilde{a}^x - a_i^x\right] - 2a^y(a_i^x - d\delta)}{2\left[(1+k)(1+c)-1\right]} - d\delta \sum_{-i \in N/\{i\}} x_{-i}$$

Compare the three optimal solutions:

- $W_i^{NC}\left(a_i^x \middle| \theta_i \in (0,1), y_i > 0\right) \geq W_i^{NC}(\theta_i = 0, y_i > 0)$ *will occur if:*

  $$\left[(1+k)(a_i^x - d\delta) - a^y\right]^2 \geq 0$$

  And concluding in:

  $$a_i^x \geq \frac{a^y}{(1+k)} + d\delta$$

  Otherwise, it will be profitable for the country to only produce clean energy.

- $W_i^{NC}\left(a_i^x \middle| \theta_i = 1, y_i > 0\right) \geq W_i^{NC}(\theta_i = 0, y_i > 0)$ *will occur if:*

  $$\left[(1+k)(a_i^x - d\delta) - a^y\right]^2 - \left[(1+k)(a_i^x - d\delta) - ((1+k)(1+c)-1)\, R^{\phi} - a^y\right]^2 \geq 0$$

  And concluding in:

  $$a_i^x \geq \frac{a^y}{(1+k)} + d\delta \qquad R \leq \left(\frac{(1+k)\left(a_i^x - d\delta\right) - a^y}{((1+k)(1+c)-1)}\right)^{\frac{1}{\phi}}$$

  Otherwise, a country will not have incentives to produce both energies, from first condition; and a country will not bind their restriction constraint, from the second condition.

- $W_i^{NC}\left(a_i^x \middle| \theta_i = 1, y_i > 0\right) \geq W_i^{NC}\left(a_i^x \middle| \theta_i \in (0,1), y_i > 0\right)$ *will occur if:*

  $$\left[(\tilde{a}^x - d\delta) - (a_i^x - d\delta)\right]^2 \leq 0$$

And concluding in:

$$R \leq \left(\frac{(1+k)\left(a_i^x - d\delta\right) - a^y}{((1+k)(1+c)-1)} \right)^{\frac{1}{\phi}}$$

From previous computations, one conclude that the optimal level x_i^{NC} *and* y_i^{NC} *are:*

$$x_i^{NC} = \begin{cases} R^{\phi}, & \textit{if } R \leq \left(\frac{(1+k)\left(a_i^x - d\delta\right) - a^y}{((1+k)(1+c)-1)} \right)^{\frac{1}{\phi}} \textit{ and } a_i^x \geq \frac{a^y}{1+k} + d\delta \\ \frac{(1+k)\left(a_i^x - d\delta\right) - a^y}{(1+k)(1+c)-1}, & \textit{if } R > \left(\frac{(1+k)\left(a_i^x - d\delta\right) - a^y}{((1+k)(1+c)-1)} \right)^{\frac{1}{\phi}} \textit{ and } a_i^x \geq \frac{a^y}{1+k} + d\delta \\ 0, & \textit{if } a_i^x < \frac{a^y}{1+k} + d\delta \end{cases}$$

And:

$$y_i^{NC} = \begin{cases} \frac{a^y - R^{\phi}}{(1+k)}, & \textit{if } R \leq \left(\frac{(1+k)\left(a_i^x - d\delta\right) - a^y}{((1+k)(1+c)-1)} \right)^{\frac{1}{\phi}} \textit{ and } a_i^x \geq \frac{a^y}{1+k} + d\delta \\ \frac{(1+c)a^y - \left(a_i^x - d\delta\right)}{(1+k)(1+c)-1}, & \textit{if } R > \left(\frac{(1+k)\left(a_i^x - d\delta\right) - a^y}{((1+k)(1+c)-1)} \right)^{\frac{1}{\phi}} \textit{ and } a_i^x \geq \frac{a^y}{1+k} + d\delta \\ \frac{a^y}{(1+k)}, & \textit{if } a_i^x < \frac{a^y}{1+k} + d\delta \end{cases}$$

Proof 7 ***STAGE 3 in the Cooperation Game***

The optimisation problem could be reduced to:

$$\begin{aligned} \max_{y_j, \theta_j} \quad & B_j\Big(\left(\theta_j R\right)^{\phi}, y_j \Big) + \beta \textstyle\sum_{j=1}^{M} y_j - C((\theta_j R)^{\phi}) - K(y_j) - D(E_N) \\ s.t. \quad & y_j \geq 0, \\ & \bar{\theta}^{IEA} \geq \theta_j \geq 0 \end{aligned}$$

Consider the cooperative maximization problem of a country j:

$$\mathcal{L}(\theta_j, y_j, \lambda) =$$

$$= a_j^x \left(\theta_j R\right)^{\phi} - \frac{\left(\theta_j R\right)^{2\phi}}{2} + a^y y_j - \frac{y_j^2}{2} - \left(\theta_j R\right)^{\phi} y_j - \left(\theta_j R\right)^{\phi} y_j + \beta \left(\sum_{j=1}^{M} y_j \right) - \frac{c \left(\theta_j R\right)^{2\phi}}{2}$$

$$-k \frac{y_j^2}{2} - d\delta \left(\sum_{i=1}^{N} x_i \right) - \lambda(\theta_j - \bar{\theta}^{IEA})$$

The Kuhn-Tucker conditions are:

$$\frac{\delta \mathcal{L}}{\delta y_j} : (a^y + \beta) - (1+k) y_j - \left(\theta_j R\right)^{\phi} \leq 0 \qquad y_j \geq 0 \tag{10}$$

$$\frac{\delta \mathcal{L}}{\delta \theta_j} : \left(a_j^x - y_j - d\delta\right) \phi \theta_j^{\phi-1} R^{\phi} - (1+c)\, \phi \theta_j^{2\phi-1} R^{2\phi} - \lambda \leq 0 \qquad \theta_j \geq 0 \tag{11}$$

$$\frac{\delta \mathcal{L}}{\delta \lambda} : \theta_j - \bar{\theta}^{IEA} \leq 0 \qquad \lambda \geq 0 \tag{12}$$

By analysing the Kuhn-Tucker conditions, we detect:

Case I: Suppose the situation where $\theta_j = 0$ and $y_j = 0$. Two results emerge from this scenario:

- *Suppose Equation 12 results in $\lambda = 0$ because $\bar{\theta}^{IEA} > 0$. Then Equation 10 results in $a^y + \beta < 0$, being not possible by construction. Additionally, Equation 11 becomes:*

  $$\frac{a_i^x - d\delta}{1+c} < (\theta_j R)^{\phi}$$

  Being not possible since $a_j^x \geq d\delta$.

- *Suppose Equation 12 results in $\lambda > 0$ because $\bar{\theta}^{IEA} = 0$. Then Equation 10 results in $a^y + \beta < 0$, being not possible by construction. Additionally, Equation 11 becomes:*

  $$\left(\phi(\theta_j)^{\phi-1} R^{\phi}\right) \left\{\left(a_j^x - d\delta\right) - (1+c)\,(\theta_j R)\right\} < \lambda$$

  Due to the Inada conditions respected by the production function:

  $$\lim_{\theta_j \to 0} \phi(\theta_j)^{\phi-1} R^{\phi} = +\infty$$

  Previous condition cannot be satisfied, since $a_j^x > d\delta$.

Case II: Suppose the situation where $\theta_j = 0$ and $y_j > 0$.Two results emerge from this scenario:

- *Suppose Equation 12 results in $\lambda = 0$ because $\bar{\theta}^{IEA} > 0$. Then Equation 10 results in $y_j = \frac{a^y+\beta}{1+k}$. Additionally, Equation 11 becomes:*

  $$\frac{\left(a_j^x - \frac{a^y+\beta}{1+k} - d\delta\right)}{(1+c)} < (\theta_j R)^{\phi}$$

  Being possible for those countries that satisfies $a_j^x < \frac{a^y+\beta}{1+k} + d\delta$ since $W_j^{IEA}(a_j^x|\theta_j = 0, y_j > 0) = \frac{1}{2}\frac{(a_y+\beta)^2}{1+k} > 0 = W_j^{IEA}(a_j^x|\theta_j = 0, y_j = 0)$.

- *Suppose Equation 12 results in* $\lambda > 0$ *because* $\bar{\theta}^{IEA} = 0$. *Then Equation 10 results in* $y_j = \frac{a^y+\beta}{1+k}$. *Additionally, Equation 11 becomes:*

  $$\left(\phi(\theta_j)^{\phi-1}R^{\phi}\right)\left\{\left(a_j^x - d\delta\right) - (1+c)\left(\theta_j R\right)\right\} < \lambda$$

  Due to the Inada conditions respected by the production function:

  $$\lim_{\theta_j \to 0} \phi(\theta_j)^{\phi-1}R^{\phi} = +\infty$$

  Being possible for those countries that satisfies $a_j^x < \frac{a^y+\beta}{1+k} + d\delta$ *since* $W_j^{IEA}(a_j^x|\theta_j = 0, y_j > 0) = \frac{1}{2}\frac{(a_y+\beta)^2}{1+k} > 0 = W_j^{IEA}(a_j^x|\theta_j = 0, y_j = 0)$.

Case III: Suppose the situation where $\theta_j > 0$ *and* $y_j = 0$. *Then, two potential scenarios emerge from Equation 12:* $0 < \theta_j < \bar{\theta}^{IEA}$ *(and* $\lambda = 0$*) or* $\theta_j = \bar{\theta}^{IEA}$ *(and* $\lambda > 0$*).*

- *Assume* $0 < \theta_j < \bar{\theta}^{IEA}$. *From Equation 11,* $\theta_j = \frac{\left(a_j^x - d\delta\right)^{\frac{1}{\phi}}}{(1+c)^{\frac{1}{\phi}}R}$. *If it is combined in Equation 10, then it will result in* $(1+c)\left(a^y+\beta\right) + d\delta < a_j^x$, *being not possible since* $a^y > a_j^x$.
- *Assume* $\theta_j = \bar{\theta}^{IEA}$. *By combining Equation 11 and 12:*

  $$\left(\phi(\bar{\theta}^{IEA})^{\phi-1}R^{\phi}\right)\left\{\left(a_j^x - d\delta\right) - (1+c)\left(\bar{\theta}^{IEA}R\right)^{\phi}\right\} > 0$$

  Being possible if $\frac{(a_j^x - d\delta)}{(1+c)} > (\bar{\theta}^C R)^{\phi}$. *Additionally, Equation 10 results in* $a^y + \beta < (\bar{\theta}^{IEA}R)^{\phi}$. *By combining both, it will end up in the following condition,* $a_j^x > d\delta + (1+c)\left(a^y+\beta\right)$, *being not possible by construction, since* $a^y > a_j^x$.

Case IV: Suppose the situation where $\theta_j > 0$ *and* $y_j > 0$. *Then, two potential scenarios emerge from Equation 12:* $0 < \theta_j < \bar{\theta}^{IEA}$ *(and* $\lambda = 0$*) or* $\theta_j = \bar{\theta}^{IEA}$ *(and* $\lambda > 0$*).*

- *Assume* $0 < \theta_j < \bar{\theta}^{IEA}$. *From Equation 11,* $\theta_j = \frac{\left(a_j^x - y_j - d\delta\right)^{\frac{1}{\phi}}}{(1+c)^{\frac{1}{\phi}}R}$. *In addition, Equation 10 shows that* $y_j = \frac{(a^y+\beta)-(\theta_j R)^{\phi}}{(1+k)}$. *When both are combined, then it will result in:*

  $$\theta_j = \frac{\left((1+k)\left(a_j^x - d\delta\right) - \left(a^y+\beta\right)\right)^{\frac{1}{\phi}}}{\left((1+k)(1+c)-1\right)^{\frac{1}{\phi}}R} \qquad y_j = \frac{(1+c)(a^y+\beta) - \left(a_j^x - d\delta\right)}{(1+k)(1+c)-1}$$

Resulting in the following welfare function:

$$W_j^{IEA}\left(a_j^x \middle| \theta_j \in \left(0, \bar{\theta}^{IEA}\right), y_j > 0\right) =$$

$$= \frac{(1+k)(a_j^x - d\delta)^2 + (1+c)(a^y+\beta)^2 - 2(a^y+\beta)(a_j^x - d\delta)}{2\left[(1+k)(1+c)-1\right]} - d\delta \sum_{-i \in N/\{j\}} x_{-i}$$

- *Assume* $\theta_j = \bar{\theta}^{IEA}$. *Also, from Equation 11,* $R \leq \left(\frac{a_j^x - y_j - d\delta}{(1+c)}\right)^{\frac{1}{\phi}} \bar{\theta}^{IEA}$. *In addition,* $y_j = \frac{(a^y+\beta) - R^\phi}{(1+k)}$. *By combining both previous conditions with Equation 9,* $R \leq \left(\frac{(1+k)\left(a_i^x - d\delta\right) - (a^y+\beta)}{((1+k)(1+c)-1)}\right)^{\frac{1}{\phi}} \bar{\theta}^{IEA}$. *Since* R *is symmetric among all countries, the level of resources will be equivalent to* $R = \left(\frac{(1+k)(\tilde{a}^x - d\delta) - (a^y+\beta)}{((1+k)(1+c)-1)}\right)^{\frac{1}{\phi}} \bar{\theta}^{IEA}$, *where* $\tilde{a}^x$ *is the minimum level of the dirty energy preference parameter such that a country will full-exhaust their resources. By using previous resource level in the optimal level of resources, we identify:*

  $$x_j = \left(\bar{\theta}^{IEA} R\right)^\phi = \frac{(1+k)\left(\tilde{a}^x - d\delta\right) - (a^y+\beta)}{((1+k)(1+c)-1)}$$

  And:

  $$y_j = \frac{(a^y+\beta) - \left(\bar{\theta}^{IEA} R\right)^\phi}{(1+k)} = \frac{(1+c)(a^y+\beta) - (\tilde{a}^x - d\delta)}{(1+k)(1+c)-1}$$

  With a welfare level equivalent to:

  $$W_j^{IEA}\left(a_j^x \middle| \theta_j = \bar{\theta}^{IEA}, y_j > 0\right) = \frac{(1+k)(a_j^x - d\delta)(\tilde{a}^x - d\delta) + (1+c)(a^y+\beta)^2}{2\left[(1+k)(1+c)-1\right]} +$$

  $$+ \frac{-(1+k)(\tilde{a}^x - d\delta)\left[\tilde{a}^x - a_j^x\right] - 2(a^y+\beta)(a_j^x - d\delta)}{2\left[(1+k)(1+c)-1\right]} - d\delta \sum_{-i \in N/\{j\}} x_{-i}$$

Compare the three optimal solutions:

- $W_j^{IEA}\left(a_j^x \middle| \theta_j \in \left(0, \bar{\theta}^{IEA}\right), y_j > 0\right) \geq W_j^{IEA}(a_j^x | \theta_j = 0, y_i > 0)$ *will occur if:*

  $$\left[(1+k)(a_j^x - d\delta) - (a^y+\beta)\right]^2 \geq 0$$

  And concluding in:

  $$a_j^x \geq \frac{(a^y+\beta)}{(1+k)} + d\delta$$

  Otherwise, it will be profitable for the country to only produce clean energy.

- $W_j^{IEA}\left(a_j^x|\theta_j = \bar{\theta}^{IEA}, y_j > 0\right) \geq W_j^{IEA}(a_j^x|\theta_j = 0, y_j > 0)$ *will occur if:*

  $$\left[(1+k)(a_j^x - d\delta) - (a^y + \beta)\right]^2 - \left[(1+k)(a_j^x - d\delta) - \left((1+k)(1+c) - 1\right) R^{\phi} - (a^y + \beta)\right]^2 \geq 0$$

  And concluding in:

  $$a_j^x \geq \frac{(a^y + \beta)}{(1+k)} + d\delta \qquad R \leq \left(\frac{(1+k)\left(a_j^x - d\delta\right) - (a^y + \beta)}{((1+k)(1+c) - 1)}\right)^{\frac{1}{\phi}}$$

  Otherwise, a country will not have incentives to produce both energies, from first condition; and a country will not bind their restriction constraint, from the second condition.

- $W_j^{IEA}\left(a_j^x\middle|\theta_j = \bar{\theta}^{IEA}, y_j > 0\right) \geq W_j^{IEA}\left(a_j^x\middle|\theta_j \in \left(0, \bar{\theta}^{IEA}\right), y_j > 0\right)$ *will occur if:*

  $$\left[(\tilde{a}^x - d\delta) - (a_j^x - d\delta)\right]^2 \leq 0$$

  And concluding in:

  $$R \leq \left(\frac{(1+k)\left(a_j^x - d\delta\right) - (a^y + \beta)}{((1+k)(1+c) - 1)}\right)^{\frac{1}{\phi}}$$

From previous computations, we conclude that:

$$x_j^{IEA}\left(\bar{\theta}^{IEA}\right) = \begin{cases} \left(\bar{\theta}^{IEA} R\right)^{\phi}, & \textit{if } R < \frac{\left(\frac{(1+k)\left(a_j^x - d\delta\right) - (a^y+\beta)}{(1+k)(1+c)-1}\right)^{\frac{1}{\phi}}}{\bar{\theta}^{IEA}} \textit{ and } a_j^x \geq \frac{a^y+\beta}{1+k} + d\delta \\ \frac{(1+k)\left(a_j^x - d\delta\right) - (a^y+\beta)}{(1+k)(1+c)-1} & \textit{if } R \geq \frac{\left(\frac{(1+k)\left(a_j^x - d\delta\right) - (a^y+\beta)}{(1+k)(1+c)-1}\right)^{\frac{1}{\phi}}}{\bar{\theta}^{IEA}} \textit{ and } a_j^x \geq \frac{a^y+\beta}{1+k} + d\delta \\ 0, & \textit{if } a_j^x < \frac{a^y+\beta}{1+k} + d\delta \end{cases}$$

And:

$$y_j^{IEA}\left(\bar{\theta}^{IEA}\right) = \begin{cases} \frac{(a^y+\beta) - \left(\bar{\theta}^{IEA} R\right)^{\phi}}{(1+k)}, & \textit{if } R < \frac{\left(\frac{(1+k)\left(a_j^x - d\delta\right) - (a^y+\beta)}{(1+k)(1+c)-1}\right)^{\frac{1}{\phi}}}{\bar{\theta}^{IEA}} \textit{ and } a_j^x \geq \frac{a^y+\beta}{1+k} + d\delta \\ \frac{(1+c)(a^y+\beta) - \left(a_j^x - d\delta\right)}{(1+k)(1+c)-1}, & \textit{if } R \geq \frac{\left(\frac{(1+k)\left(a_j^x - d\delta\right) - (a^y+\beta)}{(1+k)(1+c)-1}\right)^{\frac{1}{\phi}}}{\bar{\theta}^{IEA}} \textit{ and } a_j^x \geq \frac{a^y+\beta}{1+k} + d\delta \\ \frac{a^y+\beta}{(1+k)}, & \textit{if } a_j^x < \frac{a^y+\beta}{1+k} + d\delta \end{cases}$$

Proof 8 *There are six different scenarios to consider under the presence of the International Environmental Agreement:*

- *Suppose a country j, characterized by* $a_j^x < \frac{a^y}{1+k} + d\delta$*, decides whether to join or not the International Environmental Agreement. Under this situation, the country will continue not extracting any amount of dirty energy. Then, the country will join the coalition since any* $\beta > 0$ *will positively affect the welfare function under cooperation,* $\frac{\partial W_j^{IEA}(a_j^x|\theta_j=0,y_j>0)}{\partial\beta} > 0$ *and:*

  $$W_j^{IEA}(a_j^x|\theta_j = 0, y_j > 0) = \frac{1}{2}\frac{(a_y+\beta)^2}{1+k} > \frac{1}{2}\frac{(a_y)^2}{1+k} = W_j^{NC}(a_j^x|\theta_j = 0, y_j > 0)$$

- *Suppose a country j, characterized by* $R \geq \left(\frac{(1+k)\left(a_j^x-d\delta\right)-a^y}{(1+k)(1+c)-1}\right)^{\frac{1}{\phi}}$ *and* $a_j^x \geq \frac{a^y}{(1+k)} + d\delta$*, decides whether to join or not the International Environmental Agreement. Under this situation, the country will extract less or no amount of dirty energy,* $x_j^{IEA}\left(\bar{\theta}^{IEA} = 1\right) < x_j^{NC}$*, and more clean energy production,* $y_j^{IEA}\left(\bar{\theta}^{IEA} = 1\right) > y_j^{NC}$*, because of the conditions of the International Environmental Agreement represented by* $\beta > 0$*. In other words:*

  $$\frac{\partial W_j^{IEA}\left(a_j^x|\theta_j \in (0,1), y_j > 0\right)}{\partial\beta} > 0$$

  Since $(1+c)(a^y+\beta)+d\delta > a_j^x$ *because of construction (*$a^y \geq a_j^x$*), implying that any increase of the pro-environmental parameter will induce a rise of the welfare benefit.*

  The International Environmental Agreement will generate two potential scenarios for countries that respect the previous condition:

  I) Imagine that such country respects $a_j^x \in \left(\frac{a^y+\beta}{1+k} + d\delta, \frac{[(1+k)(1+c)-1]R^\phi+a^y}{1+k} + d\delta\right)$*, when* $\beta < [(1+k)(1+c)-1]\,R^\phi$*. Since* $(1+k)(a_j^x - d\delta) - a^y > \beta > 0$*, it will induce a higher production of clean energy in detriment of the dirty ones. As the International Environmental Agreement does not stop the production of dirty energy (because* $a_j^x > \frac{a^y+\beta}{1+k} + d\delta$*), the reduced extraction of dirty energy, in substitution of the clean energy ones, will increase the profit function.*

  $$W_j^{IEA}\left(a_j^x\middle|\theta_j \in (0, \bar{\theta}^{IEA} = 1), y_j > 0\right) > \lim_{\beta\to 0} W_j^{IEA}\left(a_j^x\middle|\theta_j \in (0, \bar{\theta}^{IEA} = 1), y_j > 0\right) =$$

  $$= W_j^{NC}\left(a_j^x\middle|\theta_j \in (0,1), y_j > 0\right)$$

  Then, those countries that satisfy previous condition, they will become signatory ones.

II) Imagine that such country respects:

$$a_j^x \in \left(\frac{a^y}{1+k}+d\delta, \min\left\{\frac{a^y+\beta}{1+k}+d\delta, \frac{\left[(1+k)(1+c)-1\right]R^{\phi}+a^y}{1+k}+d\delta\right\}\right]$$

Since $\beta > (1+k)(a_j^x-d\delta)-a^y > 0$, the implication of a sufficiently larger pro-environmental parameter could stop the extraction of dirty energy. As proved in Proof 7, the partial exhaustion is not an optimal strategy when it is compared to the no extraction of dirty energy under previous condition. Then:

$$W_j^{IEA}\left(a_j^x|\theta_j=0,y_j>0\right) > W_j^{IEA}\left(a_j^x|\theta_j\in\left(0,\bar{\theta}^{IEA}=1\right),y_j>0\right) >$$

$$> W_i^{NC}\left(a_i^x\middle|\theta_i\in(0,1),y_i>0\right)$$

Then, those countries that satisfy previous condition, they will become signatory ones.

In conclusion, those countries that satisfy $R \geq \left(\frac{(1+k)\left(a_j^x-d\delta\right)-a^y}{(1+k)(1+c)-1}\right)^{\frac{1}{\phi}}$ and $a_j^x \geq \frac{a^y}{1+k}+d\delta$ will become signatory ones.

- Suppose a country j characterized, by $R < \left(\frac{(1+k)\left(a_j^x-d\delta\right)-a^y}{(1+k)(1+c)-1}\right)^{\frac{1}{\phi}}$ and $a_j^x \geq \frac{a^y}{1+k}+d\delta$, whether to join or not the International Environmental Agreement. Under this situation, the country could extract less dirty energy or, even, stop their exhaustion, $x_j^{IEA}\left(\bar{\theta}^{IEA}=1\right) \leq x_j^{NC}$; inducing an increase in the production of clean energy, $y_j^{IEA}\left(\bar{\theta}^{IEA}=1\right) > y_j^{NC}$, because of the conditions of the International Environmental Agreement. In other words:

$$\frac{\partial W_j^{IEA}\left(a_j^x|\theta_j=\bar{\theta}^{IEA}=1,y_j>0\right)}{\partial\beta} > 0$$

Since $(1+c)(a^y+\beta)+d\delta > a_j^x$ because of construction ($a^y \geq a_j^x$), implying that any increase of the pro-environmental parameter will induce a rise of the welfare benefit.

The International Environmental Agreement will generate three potential scenarios for countries that respect the previous condition:

I) Imagine that such country respects $a_j^x \geq \frac{[(1+k)(1+c)-1]R^{\phi}+(a^y+\beta)}{1+k}+d\delta$. For any $(1+k)\left[a^y+d\delta\right]-\left[a^y+[(1+k)(1+c)-1]R^{\phi}\right] > \beta$, the country will not stop fully exhausting their dirty energy resources. According to the implication of

β in the welfare function:

$$W_j^{IEA}\left(a_j^x|\theta_j = \bar{\theta}^{IEA} = 1, y_j > 0\right) > \lim_{\beta \to 0} W_j^{IEA}\left(a_j^x|\theta_j = \bar{\theta}^{IEA} = 1, y_j > 0\right) =$$

$$W_j^{NC}\left(a_j^x|\theta_j = 1, y_j > 0\right)$$

Any country ranged in the previous condition will participate in the International Environmental Agreement.

II) *Imagine that such country respects:*

$$a_j^x \in \Big[\max\left\{\frac{a^y+\beta}{1+k} + d\delta, \frac{\left[(1+k)(1+c)-1\right]R^{\phi} + a^y}{1+k} + d\delta\right\}, ...$$

$$..., \frac{\left[(1+k)(1+c)-1\right]R^{\phi} + (a^y+\beta)}{1+k} + d\delta\Big)$$

The country will reduce the extraction level of resources, partially exhausting the current ones. Since the country respect $R < \left(\frac{(1+k)\left(a_j^x - d\delta\right) - a^y}{(1+k)(1+c)-1}\right)^{\frac{1}{\phi}}$, then:

$$W_j^{IEA}\left(a_j^x|\theta_j \in (0, \bar{\theta}^{IEA} = 1), y_j > 0\right) > W_j^{IEA}\left(a_j^x|\theta_j = \bar{\theta}^{IEA} = 1, y_j > 0\right) >$$

$$> W_j^{NC}\left(a_j^x|\theta_j = 1, y_j > 0\right)$$

In other words, any country ranged in the previous specification will partially extract the dirty energy as a response of the International Environmental Agreement.

III) *Imagine the extremal case that such country respects:*

$$a_j^x \in \left(\frac{\left[(1+k)(1+c)-1\right]R^{\phi} + a^y}{1+k} + d\delta, \frac{a^y+\beta}{1+k} + d\delta\right)$$

When $\beta > \left[(1+k)(1+c)-1\right]R^{\phi}$. Since $(1+k)(a_j^x - d\delta) - a^y < \beta > 0$, it will induce a higher production of clean energy in detriment of the dirty ones, that will stop their extraction. Under this scenario, a country that would fully exhaust before, it will completely stop the extraction of dirty energy after joining the International Environmental Agreement:

$$W_j^{IEA}\left(a_j^x|\theta_j = 0, y_j > 0\right) > W_j^{IEA}\left(a_j^x|\theta_j = \bar{\theta}^{IEA} = 1, y_j > 0\right) >$$

$$> W_j^{NC}\left(a_j^x|\theta_j = 1, y_j > 0\right)$$

Then, a country, that satisfies previous characteristics, will become signatory *even if they stop the extraction of dirty energy.*

Since all scenarios allow countries to become signatory ones, then all countries will join the International Environmental Agreement when $\bar{\theta}^{IEA} = 1$.

Proof 9 *There are three different scenarios to consider under the presence of the International Environmental Agreement:*

- *Suppose a country* j *characterized by* $a_j^x < \frac{a^y}{1+k} + d\delta$ *decides either joining or not the International Environmental Agreement. Under this situation, the country is not forced to stop their extraction of dirty energy, since it was not non-cooperatively optimal. As we have mentioned in previous proof,* $\frac{\delta W_j^{IEA}(a_j^x|\theta_j=0,y_j>0)}{\delta\beta} > 0$. *Then:*

  $$W_j^{IEA}\left(a_j^x|\theta_j = 0, y_j > 0\right) > \lim_{\beta\to 0} W_j^{IEA}\left(a_j^x|\theta_j = 0, y_j > 0\right) =$$

  $$= W_j^{NC}\left(a_j^x|\theta_j = 0, y_j > 0\right)$$

  The phase-out will not restrict those countries to join the International Environmental Agreement.

- *Suppose a country* j *characterized by* $R < \left(\frac{(1+k)\left(a_i^x - d\delta\right) - a^y}{(1+k)(1+c)-1}\right)^{\frac{1}{\phi}}$ *and* $a_i^x \geq \frac{a^y}{1+k} + d\delta$ *decides either joining or not the International Environmental Agreement. Under this situation, a country decides halting their extraction of dirty energy if:*

  $$W_j^{IEA}\left(a_j^x|\theta_j = 0, y_j > 0\right) \geq W_i^{NC}\left(a_i^x|\theta_j \in (0,1), y_j > 0\right)$$

  It will be satisfied if:

  $$\beta \geq \sqrt{\frac{(1+k)^2(a_i^x - d\delta)^2 + (1+k)(1+c)(a^y)^2 - 2a^y(1+k)(a_i^x - d\delta)}{(1+k)(1+c)-1}} - a^y > 0$$

  Notice that previous condition cannot be generalized since the threshold is positive because $a_i^x \geq \frac{a^y}{1+k} + d\delta$. *Then, depending on the size of* β, *some countries characterized by previous condition will join the International Environmental Agreement.*

- *Suppose a country* j *characterized by* $R \geq \left(\frac{(1+k)\left(a_i^x - d\delta\right) - a^y}{(1+k)(1+c)-1}\right)^{\frac{1}{\phi}}$ *and* $a_i^x \geq \frac{a^y}{1+k} + d\delta$ *decides either joining or not the International Environmental Agreement. Under* *this situation, a country decides halting their extraction of dirty energy if:*

$$W_j^{IEA}\left(a_j^x|\theta_j = 0, y_j > 0\right) \geq W_i^{NC}\left(a_i^x|\theta_j = 1, y_j > 0\right)$$

It will be satisfied if:

$$\beta \geq$$

$$\sqrt{(1+k)\frac{[(1+k)(a_i^x - d\delta)(\tilde{a}^x - d\delta) + (1+c)(a^y)^2 - (1+k)(\tilde{a}^x - d\delta)\left[\tilde{a}^x - a_i^x\right] - 2a^y(a_i^x - d\delta)]}{[(1+k)(1+c)-1]}} - a^y > 0$$

Notice that previous condition cannot be generalized since the threshold is positive because $a_i^x \geq \frac{a^y}{1+k} + d\delta$ *and* $\tilde{a}^x \geq \frac{a^y}{1+k} + d\delta$, *where* $R = \left(\frac{(1+k)(\tilde{a}^x - d\delta) - a^y}{((1+k)(1+c)-1)}\right)^{\frac{1}{\phi}}$ *and* $a_i^x \geq \tilde{a}^x$. *Then, depending on the size of* β, *some countries characterized by previous condition will join the International Environmental Agreement.*

Furthermore, it is straightforward to show that countries that full exhaust their resources before, requires a higher pro-environmental parameter than those that partial exhaust them. It is due to the fact that the former satisfies $a_i^x \geq \tilde{a}^x$, *whereas the latter satisfies* $a_i^x < \tilde{a}^x$.

Proof 10 *From the optimisation problem:*

$$\begin{aligned} \min_{\theta} \quad & d\delta N\left[\textstyle\int_{\mathbf{a^x}}^{a^y} x_i^{NC}\frac{1}{a^y - d\delta}da_i^x + \int_{\tilde{a}^x}^{\mathbf{a^x}} x_i^{IEA}(\theta)\frac{1}{a^y - d\delta}da_i^x + \int_{a^y}^{\tilde{a}^x} x_i^{IEA}(\theta)\frac{1}{a^y - d\delta}da_i^x\right] \\ s.t. \quad & \theta \in (0,1) \\ & W_j^{IEA}\left(a_j^x|\theta \in (0,1), y_j > 0\right) \geq W_j^{NC}\left(a_j^x|\theta_j \in [0,1], y_j > 0\right), \quad \forall j \in M \end{aligned}$$

By binding the participation constraint of the signatory country with the highest dirty energy characterization level, that identifies $\mathbf{a^x}$, *the* $\mathcal{L}$*agrangian function becomes:*

$$\mathcal{L}(\theta) = d\delta N\left[\int_{\mathbf{a^x}}^{a^y} x_i^{NC}\frac{1}{a^y - d\delta}da_i^x + \int_{\tilde{a}^x}^{\mathbf{a^x}} x_i^{IEA}(\theta)\frac{1}{a^y - d\delta}da_i^x + \int_{a^y}^{\tilde{a}^x} x_i^{IEA}(\theta)\frac{1}{a^y - d\delta}da_i^x\right]$$

Or:

$$\mathcal{L}(\theta) = d\delta\left[\int_{\mathbf{a^x}}^{a^y} x_i^{NC}\frac{1}{a^y - d\delta}da_i^x + (\theta R)^{\phi}\frac{\mathbf{a^x} - \tilde{a}^x}{a^y - d\delta} + \frac{[(1+k)(1+c)-1]}{2(1+k)\left[a^y - d\delta\right]}(\theta R)^{2\phi}\right]$$

By taking the first order conditions:

$$\frac{\partial \mathcal{L}}{\partial \theta} : -\frac{x_i^{NC}(\mathbf{a^x})}{a^y - d\delta}\frac{\partial \mathbf{a^x}}{\partial \theta} + \phi(\theta R)^{\phi-1}R\frac{\mathbf{a^x} - \tilde{a}^x}{a^y - d\delta} + \frac{\theta R^{\phi}}{a^y - d\delta}\left[\frac{\partial \mathbf{a^x}}{\partial \theta} - \frac{\partial \tilde{a}^x}{\partial \theta}\right] + \phi R\frac{[(1+k)(1+c)-1]}{(1+k)\left[a^y - d\delta\right]}(\theta R)^{2\phi-1} = 0$$

That it could be reduced in the following way:

$$\frac{x_i^{NC}(\mathbf{a^x})-(\theta R)^{\phi}}{a^y-d\delta}\frac{\partial \mathbf{a^x}}{\partial \theta}=\phi(\theta R)^{\phi-1}R\frac{\mathbf{a^x}-\tilde{a}^x}{a^y-d\delta}$$

Since $\frac{\partial \tilde{a}^x}{\partial \theta}=\phi R\frac{[(1+k)(1+c)-1]}{(1+k)}(\theta R)^{\phi-1}$. *Or:*

$$\frac{x_i^{NC}(\mathbf{a^x})-(\theta R)^{\phi}}{(\theta R)^{\phi}}\frac{\partial \mathbf{a^x}}{\partial \theta}=\frac{\phi}{\theta}(\mathbf{a^x}-\tilde{a}^x)$$

$$\frac{1}{\phi}\frac{x_i^{NC}(\mathbf{a^x})-(\theta R)^{\phi}}{(\theta R)^{\phi}}=\frac{(\mathbf{a^x}-\tilde{a}^x)}{\theta}\left(\frac{\partial \mathbf{a^x}}{\partial \theta}\right)^{-1}$$

$$\frac{1}{\phi}\frac{x_i^{NC}(\mathbf{a^x})-(\theta R)^{\phi}}{(\theta R)^{\phi}}=\frac{1}{\varepsilon_{\vartheta}}$$

where, in the left side, it appears the magnified (by $\frac{1}{\phi}$*) variation of the country that is dropped from the International Environmental Agreement* $\left(\frac{x_i^{NC}(\mathbf{a^x})-(\theta R)^{\phi}}{(\theta R)^{\phi}}\right)$*; and, in the right side, it shows the inverse of the elasticity of countries that full exhaust their resources at the requirements of the International Environmental Agreement at* θ*, represented with* $\varepsilon_{\theta}=\frac{\theta}{(\mathbf{a^x}-\tilde{a}^x)}\frac{\partial \mathbf{a^x}}{\partial \theta}$.

Notice that this condition works in the following way. $\frac{1}{\phi}\frac{x_i^{NC}(\mathbf{a^x})-(\vartheta R)^{\phi}}{(\vartheta R)^{\phi}}>\frac{1}{\varepsilon_{\vartheta}}$ *shows that the potential growth of aggregate emissions level out of the coalition exceeds the potential reduction generated from countries that full exhaust their resources. Then, there are incentives to increase* θ *to attract countries out of the coalition to become signatory and reduce their emissions. On the other side, when* $\frac{1}{\phi}\frac{x_i^{NC}(\mathbf{a^x})-(\theta R)^{\phi}}{(\theta R)^{\phi}}<\frac{1}{\varepsilon_{\theta}}$*, it shows that the increase in emissions generated by the country that rejects to be part of the International Environmental Agreement is less than the reduction of emissions from those signatory countries that full exhaust their resources. Then, there are incentives to still reduce* θ.

Now, consider θ^* *as the minimum cap percentage of available resources to extract that allows the grand coalition to join the International Environmental Agreement. Then:*

- *Suppose any* $\theta>\theta^*$*. No country will join the coalition since all of them are member, being* ε_{θ} *inelastic. Furthermore,* θ *will not cause any rejection of the International Environmental Agreement, inducing* $\frac{1}{\phi}\frac{x_i^{NC}(\mathbf{a^x})-(\theta R)^{\phi}}{(\theta R)^{\phi}}=0$*. Then, since:*

  $$\frac{1}{\phi}\frac{x_i^{NC}(\mathbf{a^x})-(\theta R)^{\phi}}{(\theta R)^{\phi}}<\frac{1}{\varepsilon_{\theta}}$$

  There is incentive to reduce θ.

- *Suppose any* $\theta<\theta^*$*. Some countries will decide whether to join or not the*

International Environmental Agreement. Notice that any marginal movement of θ *will drop a country from participating in such cooperation, where* $\varepsilon_\theta > 1$*. Additionally, any rejection will cause a positive increase of the current emissions, inducing* $\frac{1}{\phi}\frac{x_i^{NC}(\mathbf{a^x})-(\theta R)^\phi}{(\theta R)^\phi} > 0$*. Then, since:*

$$\frac{1}{\phi}\frac{x_i^{NC}(\mathbf{a^x})-(\theta R)^\phi}{(\theta R)^\phi} > \frac{1}{\varepsilon_\theta}$$

There is incentive to increase θ*.*

Therefore, at θ^**, there is no incentive to move since no countries will join the coalition without motivating any positive variation of the emissions. Then,* θ^* *will be the optimal maximum percentage of available resources to extract under Phase Down scenario.*

Bibliography

Andreoni, J. (1989). Giving with impure altruism: Applications to charity and ricardian equivalence. *Journal of Political Economy*, 97(6):1447–1458.

Andreoni, J. (1990). Impure altruism and donations to public goods: A theory of warm-glow giving. *The Economic Journal*, 100(401):464–477.

Bakker, B. N., Rooduijn, M., and Schumacher, G. (2016). The psychological roots of populist voting: Evidence from the united states, the netherlands and germany. *European Journal of Political Research*, 55(2):302–320.

Banerjee, A. V. (1997). A Theory of Misgovernance*. *The Quarterly Journal of Economics*, 112(4):1289–1332.

Barrett, S. (1994). Self-enforcing international environmental agreements. *Oxford Economic Papers*, 46(Supplement 1):878–894.

Barrett, S. (2006). Climate treaties and "breakthrough" technologies. *The American Economic Review*, 96(2):22–25.

Bernheim, B. D. and Whinston, M. D. (1986). Menu auctions, resource allocation, and economic influence. *The quarterly journal of economics*, 101(1):1–31.

Bhattacharya, R., Devinney, T., and Pillutla, M. (1998). A formal model of trust based on outcomes. *Academy of Management Review*, 23.

Bohm, P. (1993). Incomplete international cooperation to reduce co2 emissions: Alternative policies. *Journal of Environmental Economics and Management*, 24(3):258–271.

Calvo Ramón, E. and Rubio, S. (2013). Dynamic models of international environmental agreements: A differential game approach. *International Review of Environmental and Resource Economics*, 6(4):289–339.

Cameron, L., Chaudhuri, A., Erkal, N., and Gangadharan, L. (2009). Propensities to engage in and punish corrupt behavior: Experimental evidence from australia, india, indonesia and singapore. *Journal of Public Economics*, 93(7):843 – 851.

Caparros, A. (2016). Bargaining and international environmental agreements. *Environmental and Resource Economics*, 65:5–31.

Carraro, C. and Siniscalco, D. (1993). Strategies for the international protection of the environment. *Journal of Public Economics*, 52(3):309–328.

Carraro, C. and Siniscalco, D. (1997). R & d cooperation and the stability of international environmental agreements. In Carraro, C., editor, *International Environmental Negotiations: Strategic Policy Issues*. Edward Elgar Publishing, Cheltenham.

Chakravorty, U., Hubert, M.-H., Moreaux, M., and Nøstbakken, L. (2011). Will biofuel mandates raise food prices? *University of Alberta, Department of Economics*, 1.

Chakravorty, U. and Krulce, D. L. (1994). Heterogeneous demand and order of resource extraction. *Econometrica: Journal of the Econometric Society*, pages 1445–1452.

Chakravorty, U., Magné, B., and Moreaux, M. (2006). A hotelling model with a ceiling on the stock of pollution. *Journal of Economic Dynamics and Control*, 30(12):2875–2904.

Chakravorty, U., Moreaux, M., and Tidball, M. (2008). Ordering the extraction of polluting nonrenewable resources. *American Economic Review*, 98(3):1128–44.

Chakravorty, U., Roumasset, J., and Tse, K. (1997). Endogenous substitution among energy resources and global warming. *Journal of Political Economy*, 105(6):1201–1234.

Chambers, P. E. and Jensen, R. A. (2002). Transboundary air pollution, environmental aid, and political uncertainty. *Journal of Environmental Economics and Management*, 43(1):93–112.

Coase, R. (1960). The problem of social cost. *Journal of Law and Economics*, pages 87–137.

Cook, K. S. and Santana, J. J. (2018). Trust and rational choice. *The Oxford handbook of social and political trust*, page 253.

Cox, M. (2003). When trust matters: Explaining differences in voter turnout. *J. Common Mkt. Stud.*, 41:757.

Dahm, M. and Porteiro, N. (2008). Informational lobbying under the shadow of political pressure. *Social Choice and Welfare*, 30(4):531–559.

D'Aspremont, C., Jacquemin, A., Gabszewicz, J. J., and Weymark, J. A. (1983). On the stability of collusive price leadership. *The Canadian Journal of Economics / Revue canadienne d'Economique*, 16(1):17–25.

De Zeeuw, A. (2008). Dynamic effects on the stability of international environmental agreements. *Journal of Environmental Economics and Management*, 55(2):163–174.

Dekel, E., Jackson, M. O., and Wolinsky, A. (2008). Vote buying: General elections. *Journal of Political Economy*, 116(2):351–380.

Di Maria, C., Lange, I., and Van der Werf, E. (2014). *10 Going Full Circle: Demand-Side Constraints to the Green Paradox*, page 225. MIT Press.

Di Maria, C., Lange, I., and van der Werf, E. (2014). Should we be worried about the green paradox? announcement effects of the acid rain program. *European Economic Review*, 69:143–162. Sustainability and Climate Change: From Theory to Pragmatic Policy.

Di Maria, C., Smulders, S., and van der Werf, E. (2012). Absolute abundance and relative scarcity: Environmental policy with implementation lags. *Ecological Economics*, 74:104–119.

Duggan, J. (2018*). Bilateral lobbying: Political influence as a first price auction.

Eichner, T. and Pethig, R. (2011). Carbon leakage, the green paradox, and perfect future markets. *International Economic Review*, 52(3):767–805.

Feddersen, T. J., Sened, I., and Wright, S. G. (1990). Rational voting and candidate entry under plurality rule. *American Journal of Political Science*, pages 1005–1016.

Finus, M. (2003). Stability and design of international environmental agreements: the case of transboundary pollution. In *The International Yearbook of Environmental and Resource Economics 2003/2004: a Survey of Current Issues*, pages 82–158. Edward Elgar Publishing Ltd.

Finus, M. and Caparros, A. (2015). Game theory and international environmental cooperation. elgar res. *Rev. Econ.*

Firestone, J., Bidwell, D., Gardner, M., and Knapp, L. (2018). Wind in the sails or choppy seas?: People-place relations, aesthetics and public support for the united states' first offshore wind project. *Energy Research & Social Science*, 40:232–243.

Gerlagh, R. (2011). Too much oil. *CESifo Economic Studies*, 57(1):79–102.

Gerlagh, R. and Liski, M. (2011). Strategic resource dependence. *Journal of Economic Theory*, 146(2):699–727.

Glanville, J. L. (2016). Why does involvement in voluntary associations promote trust? examining the role of network diversity. *Sociological Inquiry*, 86(1):29–50.

Glanville, J. L. and Paxton, P. (2007). How do we learn to trust? a confirmatory tetrad analysis of the sources of generalized trust. *Social Psychology Quarterly*, 70(3):230–242.

Golosov, M., Hassler, J., Krusell, P., and Tsyvinski, A. (2014). Optimal taxes on fossil fuel in general equilibrium. *Econometrica*, 82(1):41–88.

Grafton, R. Q., Kompas, T., and Van Long, N. (2012). Substitution between biofuels and fossil fuels: is there a green paradox? *Journal of Environmental Economics and Management*, 64(3):328–341.

Grönlund, K. and Setälä, M. (2007). Political trust, satisfaction and voter turnout. *Comparative European Politics*, 5(4):400–422.

Grossman, G. M. and Helpman, E. (1994). Protection for sale. *The American Economic Review*, 84(4):833–850.

Grossman, G. M. and Helpman, E. (1996). Electoral Competition and Special Interest Politics. *The Review of Economic Studies*, 63(2):265–286.

Hagen, A. and Schneider, J. (2021). Trade sanctions and the stability of climate coalitions. *Journal of Environmental Economics and Management*, 109:102504.

Harstad, B. (2012). Buy coal! a case for supply-side environmental policy. *Journal of Political Economy*, 120(1):77–115.

Helm, C. (2000). *Economic Theories of International Environmental Cooperation*. Edward Elgar Publishing.

Hetherington, M. J. and Husser, J. A. (2012). How trust matters: The changing political relevance of political trust. *American Journal of Political Science*, 56(2):312–325.

Hoel, M. (1994). Efficient climate policy in the presence of free riders. *Journal of Environmental Economics and Management*, 27(3):259–274.

Hoel, M. (2009). Bush meets hotelling: effects of improved renewable energy technology on greenhouse gas emissions. *CESifo Working Paper Series*.

Hoel, M. (2011a). The green paradox and greenhouse gas reducing investments. *International Review of Environmental and Resource Economics*, 5(4):353–379.

Hoel, M. (2011b). The supply side of co2 with country heterogeneity*. *The Scandinavian Journal of Economics*, 113(4):846–865.

Hoel, M. (2012). Carbon taxes and the green paradox. *Climate change and common sense: Essays in honour of Tom Schelling*, pages 203–24.

Hoel, M. and Kverndokk, S. (1996). Depletion of fossil fuels and the impacts of global warming. *Resource and Energy Economics*, 18(2):115–136.

Hoffert, M. I., Caldeira, K., Benford, G., Criswell, D. R., Green, C., Herzog, H., Jain, A. K., Kheshgi, H. S., Lackner, K. S., Lewis, J. S., Lightfoot, H. D., Manheimer, W., Mankins, J. C., Mauel, M. E., Perkins, L. J., Schlesinger, M. E., Volk, T., and Wigley, T. M. L. (2002). Advanced technology paths to global climate stability: Energy for a greenhouse planet. *Science*, 298(5595):981–987.

Hooghe, M. and Marien, S. (2013). A comparative analysis of the relation between political trust and forms of political participation in europe. *European Societies*, 15(1):131–152.

Hudson, J. (2006). Institutional trust and subjective well-being across the eu. *Kyklos*, 59(1):43–62.

Jus, D. and Meier, V. (2015). Announcing is bad, delaying is worse: Another pitfall in well-intended climate policy. *Jahrbücher für Nationalökonomie und Statistik*, 235(3):286–297.

Kahneman, D. and Knetsch, J. L. (1992). Valuing public goods: The purchase of moral satisfaction. *Journal of Environmental Economics and Management*, 22(1):57–70.

Karlstrøm, H. and Ryghaug, M. (2014). Public attitudes towards renewable energy technologies in norway. the role of party preferences. *Energy Policy*, 67:656–663.

Kirchsteiger, G. and Prat, A. (2001). Inefficient equilibria in lobbying. *Journal of Public Economics*, 82(3):349–375.

Laffont, J.-J. (1999). Political economy, information and incentives. *European Economic Review*, 43(4-6):649–669.

Larsen, C. A. (2007). How welfare regimes generate and erode social capital: The impact of underclass phenomena. *Comparative Politics*, 40(1):83–101.

Lewicki, R. J. and Bunker, B. B. (1995). Trust in relationships. *Administrative Science Quarterly*, 5(1):583–601.

Lewis, J. D. and Weigert, A. (1985). Trust as a social reality. *Social forces*, 63(4):967–985.

Mocan, N. (2008). What determines corruption? international evidence from microdata. *Economic Inquiry*, 46(4):493–510.

Murdoch, J. C. and Sandler, T. (1997a). Voluntary cutbacks and pretreaty behavior: the helsinki protocol and sulfur emissions. *Public Finance Review*, 25(2):139–162.

Murdoch, J. C. and Sandler, T. (1997b). The voluntary provision of a pure public good: The case of reduced cfc emissions and the montreal protocol. *Journal of Public Economics*, 63(3):331–349.

Nordhaus, W. (2015). Climate clubs: Overcoming free-riding in international climate policy. *American Economic Review*, 105(4):1339–70.

Obydenkova, A. V. and Arpino, B. (2018). Corruption and trust in the european union and national institutions: Changes over the great recession across european states. *JCMS: Journal of Common Market Studies*, 56(3):594–611.

Palfrey, T. R. (1984). Spatial equilibrium with entry. *The Review of Economic Studies*, 51(1):139–156.

Proschan, F. (1965). Peakedness of distributions of convex combinations. *The Annals of Mathematical Statistics*, 36(6):1703–1706.

Ramstetter, L. and Habersack, F. (2020). Do women make a difference? analysing environmental attitudes and actions of members of the european parliament. *Environmental Politics*, 29(6):1063–1084.

Rohrschneider, R. (2002). The democracy deficit and mass support for an eu-wide government. *American Journal of Political Science*, 46(2):463–475.

Rubio, S. J. (2017). Sharing R&D investments in breakthrough technologies to control climate change. *Oxford Economic Papers*, 69(2):496–521.

Santos, A. S., de Abreu, V. H. S., de Assis, T. F., Ribeiro, S. K., and Ribeiro, G. M. (2021). *An Overview on Costs of Shifting to Sustainable Road Transport: A Challenge for Cities Worldwide*, pages 93–121. Springer Singapore, Singapore.

Schultz, C. (1996). Polarization and inefficient policies. *The Review of Economic Studies*, 63(2):331–343.

Sengers, F., Raven, R., and Van Venrooij, A. (2010). From riches to rags: Biofuels, media discourses, and resistance to sustainable energy technologies. *Energy Policy*, 38(9):5013–5027. Special Section on Carbon Emissions and Carbon Management in Cities with Regular Papers.

Serra, D. (2006). Empirical determinants of corruption: A sensitivity analysis. *Public Choice*, 126(1/2):225–256.

Sinclair, P. J. N. (1992). High Does Nothing and Rising Is Worse: Carbon Taxes Should Keep Declining to Cut Harmful Emissions. *The Manchester School of Economic & Social Studies*, 60(1):41–52.

Sinclair, P. J. N. (1994). On the optimum trend of fossil fuel taxation. *Oxford Economic Papers*, 46:869–877.

Sinn, H.-W. (2008a). Das grüne paradoxon: Warum man das angebot bei der klimapolitik nicht vergessen darf. *Perspektiven der Wirtschaftspolitik*, 9(s1):109–142.

Sinn, H.-W. (2008b). Public policies against global warming: a supply side approach. *International tax and public finance*, 15(4):360–394.

Sinn, H.-W. (2012). *The green paradox: a supply-side approach to global warming*. MIT press.

Smulders, S., Tsur, Y., and Zemel, A. (2012). Announcing climate policy: Can a green paradox arise without scarcity? *Journal of Environmental Economics and Management*, 64(3):364–376. 2010 Monte Verita Conference on Sustainable Resource Use and Economic Dynamics(SURED).

Strand, J. (2007). Technology treaties and fossil-fuels extraction. *The Energy Journal*, 28:129–141.

Strazzera, E. and Statzu, V. (2017). Fostering photovoltaic technologies in mediterranean cities: Consumers' demand and social acceptance. *Renewable Energy*, 102:361–371.

Svensson, J. (2003). Who must pay bribes and how much? evidence from a cross section of firms. *The Quarterly Journal of Economics*, 118(1):207–230.

Tabi, A. and Wüstenhagen, R. (2017). Keep it local and fish-friendly: Social acceptance of hydropower projects in switzerland. *Renewable and Sustainable Energy Reviews*, 68:763–773.

Tahvonen, O. (1995). Dynamics of pollution control when damage is sensitive to the rate of pollution accumulation. *Environmental and Resource Economics*, 5(1):9–27.

Tahvonen, O. (1997). Fossil fuels, stock externalities, and backstop technology. *The Canadian Journal of Economics / Revue canadienne d'Economique*, 30(4a):855–874.

Taufik, D. and Venhoeven, L. (2018). *Emotions and Pro-Environmental Behaviour*, chapter 19, pages 189–197. John Wiley & Sons, Ltd.

Tirole, J. (1996). A Theory of Collective Reputations (with applications to the persistence of corruption and to firm quality). *The Review of Economic Studies*, 63(1):1–22.

Tranter, B. (2011). Political divisions over climate change and environmental issues in australia. *Environmental Politics*, 20(1):78–96.

Treisman, D. (2000). The causes of corruption: a cross-national study. *Journal of Public Economics*, 76(3):399 – 457.

Ulph, A. and Ulph, D. (1994). The optimal time path of a carbon tax. *Oxford Economic Papers*, 46:857–868.

Uslaner, E. M. (2002). *The moral foundations of trust.* Cambridge University Press.

van der Ploeg, F. and Withagen, C. (2012). Is there really a green paradox? *Journal of Environmental Economics and Management*, 64(3):342–363. 2010 Monte Verita Conference on Sustainable Resource Use and Economic Dynamics(SURED).

Van Erkel, P. F. and Van der Meer, T. W. (2016). Macroeconomic performance, political trust and the great recession: A multilevel analysis of the effects of within-country fluctuations in macroeconomic performance on political trust in 15 eu countries, 1999–2011. *European Journal of Political Research*, 55(1):177–197.

Van Long, N. and Stähler, F. (2014). Resource extraction and backstop technologies in general equilibrium. *Climate policy and nonrenewable resources: The green paradox and beyond*, page 151.