**Multi-stakeholder coalitions, subjective games, and equitable equilibria in the Great Bear Rainforest agreement.**

**Abstract**

Forest governance is often at the root of multiple-level conflicts among actors with various forms of power that involve a diverse array of costs and risks and negative impacts in terms of both efficiency and justice. However, if properly managed, conflicts can be a positive force for institutional change and innovation whereby several cross-scale interests and objectives can be coordinated and new social and power relationships be forged.

This study uses game theory modelling to investigate the processes leading to the historical agreement over the Great Bear Rainforest after decades of “War in the woods” contestation among Provincial Government, First Nations, environmental NGOs, and timber companies and partial compliance with it. We move from the standard game theory approach by applying the subjective games model to the ex-ante phase – the transformation of conflict into dialogue for an agreement through changes that led agents to reframe their subjective models from a social dilemma to a stag hunt. Then, we apply the psychological games model to study the ex-post stage, that is implementation of a more environmentally and socially just agreement. The overall aim is to study (i) the formation of conditions that led to the multi-stakeholders agreement and (ii) the conditions under which compliance is feasible. The approach is consistent with a view of environmental policy as resulting from interactions between organised various decision-makers, interest groups, and coalitions that shift bargaining power.

1. **Introduction**

Forests are multifunctional resources that provide a wide range of goods and services, from the local to the global scale (Frischmann, 2012; Geores, 2003; Gong, 2002; Oldekop et al., 2020). Because of this, forests governance and management are often at the root of multiple-level conflicts (Eckerberg & Sandström, 2013; Hellström, 2001; Mola-Yudego & Gritten, 2010; Nouisiainen & Mola-Yudego, 2022; Yasmi, Kelley and Enters, 2012). The Food and Agriculture Organization (FAO) of the United Nations characterizes natural resources conflicts as “disagreements and disputes over access to, and control and use of, natural resources” (Matiru, 2000:1). They arise when users have competing demands or different management priorities, and those of some groups are excluded from policies, programs and projects (Matiru, 2000). Involved parties might also have the same claims but different views on how the resource should be distributed to meet their needs, or on exploitation and distribution procedures (White et al., 2009), as well as different degrees of power to influence decision-making procedures and outcomes (Morrison et al., 2019; Kashwan et al., 2021). Different and competing aspirations for using and managing the same resource are at stake (Adams et al., 2003; Matilainen et al., 2017), including immaterial and immeasurable higher-level principles and core values, as well as technical and evidence-based arguments (Sranko, 2011).

Whatever their nature, scale, geographical scope and time coverage, conflicts imply an array of costs and risks, to the point that they can be described as “a lose-lose situation and a manifestation of governance failure” (Eckerberg & Sandström, 2013:1). The result is a negative impact on both efficiency and fairness: indeed, beyond being at the roots of forests deterioration and depletion, conflicts may result in unequal access to or abuse of the good by one group on behalf of another. These processes, in turn, can further exacerbate power and wellbeing asymmetries (Chiappero, Von Jacobi and Fabbri, 2015). On the other hand, if properly handled, conflicts can give rise to innovative institutional solutions, new collaborative governance arrangements and progressive policy reforms (Eckerberg & Sandström, 2013; Hafner et al., 2003; Hellström, 2001; Nouisiainen & Mola-Yudego, 2022). Since the root of these issues usually lays in decision-making problems over joint production of several concurring Ecosystem Services (Nelson et al., 2009), transforming disputes into agreement is more about coordinating several cross-scale interests and objectives (Berkes, 2002) rather than optimizing a single objective function.

In this research, we develop a game-theory (GT) model to investigate the prospects for innovative forest governance agreements and conformity with them, which may come from cross-scale (Berkes, 2002) institutional interactions.

Analyzing multi-stakeholder decision-making processes on how to use a natural resource from a GT viewpoint appears particularly illuminating (Parrachino, Zara and Patrone, 2006). Indeed, GT investigates the mathematical modelling of strategic interactions among decision makers, particularly when one player's actions affect others (Enders, 2004; Parrachino et al., 2006; Parrachino et al., 2006; Zara et al., 2006). A game is defined as any situation involving at least two players who recognize that their payoffs are influenced not only by their decisions but also by the decisions of other players (Hargreaves Heap et al., 1992). GT also proposes that games representing social situations are essentially mixed-interest games (Harsanyi, 1977), with a common interest in avoiding a mutually destructive non-agreement outcome and a conflict of interest in deciding which of the many coordination solutions to agree on. Thus, while cooperative tactics may produce better results, there are various obstacles to achieving them, including a lack of enforcement mechanisms and equilibrium property, social dilemmas (Zara et al., 2006), coordination problems and multiplicity of equilibrium solutions. Multi-stakeholders’ models for natural resources governance are particularly fascinating in this regard since they frequently disclose critical insights on such challenges as GT normally analyses.

This study uses GT mathematical modeling to explore the processes that led to the historical agreement and final compliance with the Great Bear Rainforest (GBR), the world's largest unspoiled temperate rainforest located in British Columbia (BC). However, we depart from the standard game theory approach. Indeed, we begin by applying the *subjective games model* (Aoki, 2001) to the *ex-ante* phase, which is the transition from a conflictual into prospects for an agreement. Then, we use *psychological games* (Cecchini Manara & Sacconi, 2019b, 2021; Geanakoplos et al., 1989; Grimalda & Sacconi, 2005, 2007; Rabin, 1993) to study the *ex-post* stage, that is compliance with the agreement.

The overarching goal is to utilize GT to investigate the formation of conditions (i) that led to the multi-stakeholder agreement on the GBR – the ex-ante game, and (ii) under which compliance is achievable - the implementation game. The approach is congruent with an understanding of environmental policy as the consequence of interactions between organized diverse decision makers and interest groups (Dijkstra BR, 1999; Endres, 2004). Moreover, the results open a discussion on the issue of green conservation v. social equity (Chen et al., 2022; Franks et al., 2018; Oldekop et al., 2015; Roe, 2008) , as – especially related to forests governance – certain “green” solutions such as the creation of Protected Areas are more and more criticized as new forms of racism and colonialism in the name of “environmental protection” (Mombeshora & Le Bel, 2009; Wilkie et al., 2006).

The next Sections are organized as follows. Section 2 presents the theoretical framework, which serves as the building block for the following discussion. Section 3 illustrates the case study, describing the history behind the transformation of a “War in the Woods” (Cashore et al., 2011) into what is celebrated as one of the most advanced agreements on forests governance and biodiversity conservation (Armstrong, 2009; IPCAs, 2018; XXXX). Section 4 introduces an analytical model of the story and a simplified representation of the factors that made the conflict transformation possible, therefore focusing on the ex-ante phase. Here, we refer to framing games and institutional complementarity (Aoki, 2001). Section 5 focuses on the ex-post conformity and presents a model of a psychological game, taking and extending the one presented by Degli Antoni & Sacconi (2013) to study social responsibility, activism and boycotting in the relationships between firms and stakeholders. In the conclusions, we outline the main findings, underline some limits of our study and give insights for future research.

**2. Theoretical framework**

In his book *Toward Comparative Institutional Analysis*, Masahiko Aoki (2001) gives a conceptualization of *institutions as equilibria*, building on a game theory approach and other interdisciplinary contributions. In his account, an institution is “a self-sustaining system of shared beliefs about a salient way in which the game is repeatedly played" (Aoki, 2001:10). Each player has a “limited capacity of cognition, evaluation and calculation (Aoki, 2001:205), by which she envisions a subjective and compressed model of the game, that is a summary depiction of it where only subsets of conceivable actions for large games are displayed. Then, in the iterative process, such an initial frame constrains or empowers the player's strategies, determining the resulting equilibrium selection (Aoki, 2001; Cecchini Manara & Sacconi, 2019a). This conclusion reinforces the initial summary representation as long as it maintains an equilibrium.

However, new information or collective action in which agents participate or are informed about as emanating from complementary choice arenas may activate a revision in the player’s frame and possibly ignite a new set of activities. If this reassessment results in mutual coherence between the various frames of individual agents, what Aoki (2001:232) refers to as a new common system of beliefs, we have the foundation for a new organization.

A new endogenous equilibrium state emerges because of strategic choices made by agents who, by acting in accordance with these new shared beliefs, go through an equilibrium selection process, eventually achieving a state in which their action combination replicates and maintains the equilibrium, i.e., a new institution. As a result, the institution is self-sustaining as long as agents believe in it as a substantial summary representation of behavior regularity, in line with a given inter-subjective mental model of the game played in the domain. When external shocks, internal domain crises, capacity growth, or a combination of these disrupt the shared system of belief, agents may seek new avenues of activity. Because the latter may exist in dimensions that were previously inactive, the agent's search for new repertoires of action may lead to the discovery of a new subjective frame. When this occurs for a large enough number of people, the survival of the old institution is jeopardized.

Cognitive representations and choices of agents playing in a certain domain can also be parametrically affected by prevailing institutions in other domains, without agents strategically coordinating their choices across them. Aoki (2001) refers to such interdependencies as *institutional complementarities*, claiming that they can drive agents to revise their perception of the game's structure and discover new ways of doing things, potentially leading to institutional innovation.

In this work, we use Aoki’s theoretical framework to explain the *ex-ante* phase, that is the transformation of a multi-stakeholders’ conflict where cooperation has no mutually beneficial outcomes into multi-stakeholders’ interactions with feasible prospects for agreement and cooperation, which the parties could exploit through some form of agreement or explicit coordination (Section 4). In this shift, a new frame emerges that allows the parties to see the possibility for solutions that could be object of an agreement and be sustained in equilibrium.

We anticipate that participants' subjective models of the game will evolve because of environmental changes, collective actions, and institutional complementarity (Aoki, 2001), resulting in a new repertory of actions, consequence functions, and hence payoffs. While players witness new non-cooperative strategies aimed at assuring a certain level of security for each agent, unilateral defection becomes less prominent. These strategies have always been there, but they are now becoming prominent as the parties do not see the game as a mutual distrust game anymore. Cooperating by agreement is a possible outcome in the updated framed game - a Stag Hunt game - because it is supported by an equilibrium. This alters the character of local stakeholders' interactions.

The process we describe through Aoki’s (2001) approach explains how, in the interaction between certain players, a specific mental model emerges that allows to see the cooperation between business organizations and ENGOs as an arena in which a mutual limitation agreement could credibly replace mutual distrust. In Section 4, we specifically enlighten how information coming from other arenas of interaction - mainly between the logging companies and their intermediate customers – later on defined as strong stakeholders Ss - along all the supply chain through ENGOs’ lobbying and market campaigning, allows a shift in the framing of the game played amongst ENGOs and E, so that they start seeing the possibility of a feasible cooperation in equilibrium, even if the prospect of disagreeing remains viable. The same is true for the interaction amongst First Nations and the Provincial Government initially characterized by distrust.

These shifts open the route to a wider agreement amongst the different parties here involved, with the Provincial Government acting the rule maker and attempting to investigate the potential for a broader agreement that includes both environmental and social provisions.

We do not, however, examine the game associated with the bargaining table. It is sufficient to show that the prospect of credible agreements has evolved, and we consider that a bargaining agreement could correspond to a fair allocation of the surplus, reflecting the Nash Bargaining Solution. What we are interested in, instead, is the *ex-post* game that follows by taking the bargaining phase as given, with the inherent risk that it results in a cheap talk game without effectiveness on the players’ compliance. Indeed, it is not to be taken for granted neither that the agreement is binding *per se* nor that the government has the capacity to enforce it. Moreover, we anticipate that when the conditions for sitting at a multi-stakeholder negotiation table are met, there is a problem in that only agreements limited to business organizations and environmental NGOs (ENGOs) are implementable in equilibrium. Thus, First Nations’ claims, which include demands for socially equitable outcomes where forest protection goes hand in hand with safeguarding their way of life with the meaning they ascribe to living in contact with the forest, would be left out from the agreement. Hence, the negotiating table for a big multi-stakeholder agreement appears doomed to either fail despite the positive shift in the framing of the game, or to finish in a sustained agreement of green collusion between business organizations and ENGOs, to the disadvantage of Indigenous communities.

In Section 5, we investigate these issues using the contribution of psychological game theory, namely the model of conformity preference and an additional piece of game theory in relation to Aoki's concept of framed games (Cecchini Manara & Sacconi, 2019b, 2021; Geanakoplos, Pearce and Stacchetti, 1989; Grimalda & Sacconi, 2005, 2007; Rabin, 1993). As a result, a 4-player game with numerous conceivable outcomes, including major environmental and social constraints, appears. At a first glance, however, the equilibrium property is only held by the outcome in which the logging companies and the environmentalists agree on a limitation of extraction without regard for the First Nations’ claims. It is no coincidence that this is the situation; indeed, it corresponds to what emerges from the change in the frame through which the game between logging companies and ENGOs is viewed. Thus, an agreement with broader content would not be sustained and the game designer function of the Provincial Government would appear doomed to fail.

However, reality shows a different story, as the agreement has gained substantial though not perfect support and compliance in the years thereafter. We explain these empirical findings by applying the theory of conformity preferences in psychological games to the case at stake. Indeed, we develop our second hypothesis by modelling the implementation game as a psychological game (Section 5). Psychological games (Geanakoplo et al., 1989; Rabin, 1993)are especially useful for understanding endogenous forces that an agreement can activate to induce compliance equilibria (Cecchini Manara & Sacconi, 2019b, 2021; Grimalda & Sacconi, 2005, 2007; Sacconi & Faillo, 2010). To model this stage, we take inspiration from and readapt Degli Antoni & Sacconi’s work (2013) on “Social responsibility, activism and boycotting in a firm–stakeholders network of games with players’ conformist preferences”.

**3. The Great Bear Rainforest Case**

The beginning of the “War in the Woods” (Cashore et al., 2001; Forests Ethics, Greenpeace and Sierra Club, 2016) dates to the 1980s, when some of the First Nations (FNs) traditionally living in the area and environmental movements and grassroots organizations (ENGOs), started protesting against logging activities carried out by BC timber industries in old-growth forests (Affolderbach, 2011: Saakiroski et al., 2013). FNs and ENGOs defended the rainforest as being of both local and global value (Armstrong, 2009) and called for increased protection and new type of forest managements (Raitio & Saakiroski, 2014). Yet, ENGOs were mostly advocating for the environmental value of the forest, as a hotspot of wildlife and biodiversity richness, to the extent that they renamed it as the “Great Bear Rainforest” (GBR) after the white-coated Kermode Bear endemic to the region. The FNs, instead, claimed that commercial logging of the area was undermining their livelihoods and socio-economic development, and it had been implemented without their consent, further undermining their self-governance (Raitio & Saakiroski, 2012). In addition, they refused pure conservationist arguments in favor of strictly protected areas, as this would have denied their rights to hunt, harvest and take ceremonial logs from the forests. (IPCAs, 2018). On their part, forestry companies strongly opposed any request of reducing timber supply, as this would have threatened the profitability of their operations and negatively affected the regional economy (Armstrong, 2009; Cashore et al., 2011).

In the middle of such disagreement over the direction of BC forest policy (Saakiroski, Raitio and Barry, 2013), the Provincial Government (ProvGov) found itself in a very uncomfortable position. On the one hand, it had to represent the public (environmental) interest, which ENGOs claimed they were defending, while on the other it benefitted economically from logging, the main economic driver of the region at the time. Moreover, it had to ensure that FNs’ interests were represented in any ultimate deal. At the beginning, the ProvGov was mostly acquiescent with the industries, but it also tried to introduce a collaborative planning system - the Land and Resource Management Planning (LRMP) table. Yet, ENGOs thought this entire process would have served only as a cover to continue logging, and FNs believed that taking part into it would have jeopardized the possibility to have their land claims satisfied. Thus, they both decided to stay out from it and pursue different strategies (Howlett et al., 2009; Raitio & Saakiroski, 2014; Wilson, 1998; Smith, 2010). Indeed, FNs decided to pursue a legal path by bringing cases to Courts, while ENGOs started a huge campaign targeting the BC forest industries and their international customers and investors (Raitio & Saakiroski, 2014). These strategies contributed to change the ProvGov and logging companies’ attitude respectively (Affolderbach, 2011; Sranko, 2011).

As the market campaigns of ENGOs had reached their international customers, the logging companies extracting timber in the areas began to worry that their reputation would be harmed, with far-reaching economic consequences. A critical point occurred in 1999, when a German delegation of papermakers and magazine publishers, after visiting active logging sites and having met some of the FNs, threatened the cancellation of contracts unless an acceptable solution to the conflict was found (Armstrong, 2009; Raitio, Saakiroski and Barry, 2013). Under this pressure, the logging industries decided to move from an “attack and defend” strategy towards a more conciliatory one (Affolderbach, 2011; Armstrong, 2009; Raitio & Saakiroski, 2014). Some of them – under the so-called Coast Forest Conservation Initiative (CFCI) - entered bilateral negotiations with ENGOs, which resulted into an agreement over basic principles for the development of joint solutions and the suspension of respective activities that were at the roots of the conflict (Armstrong, 2009). Through the so-called Joint Solutions Project (JSP), logging companies and ENGOs agreed upon a set of principles for joint solutions – precautionary principle, protection of habitat for fish and wildlife, harvesting techniques with low environmental impact and high timber value among others (Armstrong, 2009) – and on the creation of the Coast Information Team (CIT), an independent team of experts that would have provided knowledge and learning all along the process (Raitio et al., 2013; Smith, 2010). Moreover, ENGOs obtained moratorium on logging activities in contested areas, to which the industries agreed in return for a halt to market campaigns (Armstrong, 2009; Raitio, Saakiroski and Barry, 2013).

On the other hand, both the FNs and the ProvGov felt that industries and ENGOs were coming to agreements on territories that were their own (UBCM, 2000). Thus, by 2001, the ProvGov began to sign protocols with the FNs so to enter a Government-to-Government (G2G) relation, by which they undertook a commitment for shared-decision making and negotiation of agreements over land use planning and resource management. This decision-making table served to counterbalance the ENGOs-industries alliance, or stakeholders table (Cullen et al., 2010; Raitio & Saakiroski, 2014).

In the meanwhile, the coastal FNs, who had until then worked mostly independently of each other to promote their rights and title, had started establishing formal coalitions such as the Coastal First Nations Turning Point Initiative (CFN). At the same time, the relation between ENGOs and FNs became stronger: the former recognized their need to take FNs’ concerns seriously while the latter realized they could benefit from market campaigns. Indeed, the campaigns contributed to put the region on the global agenda and created the visibility and pressure FNs needed to have more influence on the ProvGov and the industries. It is against this background that a critical meeting took place in 2001, where the FNs expressed their support for environmental pressure but asked the ENGOs to recognize their rights and title over those territories and consider FNs’ well-being in their claims (Raitio & Saakiroski, 2014; Raitio et al., 2013). In other words, FNs wanted the issue of equity to be brought into the agreement: they claimed that colonization had already had a strong impact on their ability to self-manage land and ENGOs were now approaching the problem with a neo-colonialist mind, thinking mostly from a strategic and pure conservationist point of view rather than from an equity perspective.

In the end, the ProvGov retook control of the entire situation and brought key stakeholders to the formal LRMP Completion Table in December 2001. As Raitio & Saakiroski (2013) well explain, the ProvGov itself was strongly dependent on forest industries for income and employment and thus wished parties could arrive at a permanent decision to be communicated to markets. On the other side, firms, ENGOs and FNs needed certainty and ProvGov’s legitimacy, leadership and expertise. Moreover, failure to reach a consensus would have induced the risk of a unilateral decision by the ProvGov itself.

The entire process culminated into BC Premier Gordon Campbell’s announcement, in February 2006, that an historical agreement had been reached over the GBR. A concrete outcome of FNs’ increased influence was the establishment of a new category of protected areas called Conservancies, which, differently from strictly protected areas that exclude any human activity, allow for the FNs’ traditional social, cultural and ceremonial uses (Howlett et al., 2009:389; IPCAs, 2018; Saakiroski & Barry, 2013; Raitio & Saakiroski, 2012). The agreement established a network of protected areas that was off-limits to industrial logging (33 per cent of the region), improved logging activities over an additional part of the area (37 per cent of the region) and increased involvement of FNs’ in decision-making and conservation financing for them (IPCAs, 2018). Three years later, in 2009, the parties agreed to a five-year plan to review the implementation of the Ecosystem Based Management (EBM) plan, with the aim of "concurrently moving to high levels of ecological integrity and high levels of human well-being and if that is not possible, to make meaningful increments to both" (BC Gov News Releases 29.1.2014). Then, in 2016, the government passed the Great Bear Rainforest Land Use Order (GBRLUO) and the Great Bear Rainforest (Forest Management) Act (BC Gov News Releases 1.3.2016 and 21.12.2016) with the objective of conserving 85% of the forest and 70% of old growth. Importantly, the agreement gave FNs a higher stake in forest management in the region and saw an increase in the allocation of forest carbon-offset credits to be sold and used for development projects of importance to them[[1]](#footnote-1). A Coast fund was also created to finance indigenous-led conservation and the development of new sustainable businesses (Armstrong, 2009; Smith, 2010).

Overall, the GBR agreement is considered to have brought a notable situation improvement (Walker & Daniels, 1996), as it contributed to (i) protect the forest through a more ecological perspective and slow down the harvesting rate and (ii) increase FNs’ rights and decision-making powers over management of land and land revenues. However, the situation is not black and white; many ENGOs have been complaining about delays over the implementation of some key protection measures such as plans for the most endangered ecosystems, which creates ecological uncertainty (Raitio et al., 2013; Rainforest Solutions Project, Press Release 12.02.2020). Hence, actual ex-post compliance with the agreement is only partial, as some logging industries have failed to disclose information on mapping and what they were going to protect (Rainforest Solution Project, 2018). The fact that, when the conflict scaled down, ENGOs lost leverage in the market and the media, is one of the main reasons for such partial failure. Indeed, international market campaigns targeting customers of logging companies harvesting in the region had been the key factor for changing the conflict (Affolderbach, 2011; Bulkeley, 2005). FNs instead, are more satisfied and recognize increased benefits, so that they continue to work in partnership with the ProvGov to defer logging of old growth, while developing a new approach to sustainable forest management. Recently, new measures have been introduced to sustain the Great Bear Rainforest under the ecosystem-based management approach, also with a view to increase FN’s role and responsibility in co-managament (BC Gov News Release 27.7.2023).

**4. From multi-stakeholders’ conflict to opportunities for agreement and cooperation: the GBR case explained through game theory**

The GBR case is an example of a complex conflict involving multiple stakeholders, which was turned into an agreement for multi-stakeholder governance particularly thanks to FNs and ENGOs’ ability to organize and mobilize at multiple scales.

This Section attempts to model the trajectory from a multi-stakeholder’s conflict to interactions in which there are feasible prospects for agreement and cooperation that might be sustained in equilibrium. We give a reconstruction separated into stages, as we regard them as distinct games played in discrete time instants along an historical process, rather than components of a single game. This allows the solution of the game in each stage not to consider the cumulative payoffs achieved in all previous phases[[2]](#footnote-2). We adopt the subjective game models approach (Aoki, 2001) to illustrate how the game models representing the historical situation change at different stages due to players cognitive disequilibrium. Changes in information from external factors and complementary games, make agents shift their own subjective representation of the game at some point; from that moment on, they believe they can play a different game with new strategic opportunities and new equilibrium solutions.

*~~4~~.1 A shift in the players’ subjective game models*

The very opening stage of the story (Phase 1), when the ProvGov tries to set up a LRMP table but both the ENGOs and FNs decide to stay out, is a phase of huge distrust, as ENGOs and FNs fear the agreement would not consider equity issues and would simply be a cover up for continuous logging. Thus, there is not a proper multi-actor bargaining table and the parties play separate games, with ENGOs and FNs negotiating with some of the logging industries and the ProvGov respectively (Raitio, Saakiroski and Barry, 2013). By assumption, at this stage the various ENGOs, FNs and logging companies are considered as three single players[[3]](#footnote-3), hereinafter ENGO, FN and E. All of them, as well as the ProvGov, are rational decision makers who maximize their own objective function.

Each of the two parallel interactions, one between ENGO and E and the other between FN and the ProvGov, can be modelled as a 3x3 game that encompasses various subgames. This is a more complex representation than what players can see and perform since, due to limited awareness, they focus on specific subgames based on their mental models and their strategic interactions. The game is presented in Figure 1 in its extensive form and in Table 1 in its normal form

D

1

2

1

c

d

S

1

N

2

c

c

d

c

c

c

c

c

b

b

d

a

a

a

c

c

N

N

S

C

D

C

C

C

D

D

D

C

**Figure 1.** The extensive form of the game depicting the relationship between ENGO-E in Phase 1. The letters in the column represent the payoffs obtained by ENGOs and E, with a > b > c > d. A game with the same structure is played between FN and the ProvGov.

|  |  |  |  |
| --- | --- | --- | --- |
| **E**  **ENGO** | **S, D** | **S, C** | **N** |
| **S, D** | (c,c) | (a,d) | (c,c) |
| **S, C** | (d,a) | (b,b) | (a,c) |
| **N** | (c,c) | (c,d) | (c,c) |

**Table 1.** The normal form of the game depicting the relationship between ENGO and E in Phase 1. The letters represent the payoffs obtained by ENGO and E, with a > b > c > d. A game with the same structure is played between FN and the ProvGov.

Figure 1 shows that at first, each participant can decide whether to

* negotiate and subscribe an agreement (S); or
* do not negotiate and act without looking for cooperation (N).

If the player plays S, then she can decide to

* comply with the agreement (C); or
* do not comply with it and take advantage of the other’s compliance if it occurs (D).

If instead the player initially plays N, she can continue by doing it herself without looking for the other’s cooperation.

For example, considering the game between E and ENGO:

* By playing S-C, E agrees on a rate of forest to be protected and be set aside from logging, as well as a new way of logging the rest of the area, and commits to abide by the agreement; correspondingly, ENGO genuinely undertake to stop any campaign against E.
* By playing S-D, the player agrees on the same points but defect and opportunistically exploit the other's cooperation if it occurs. For E, this means continuing logging even though ENGO has renounced market campaigns, while for ENGO it means continuing market campaigns against E, even against those who comply.
* By playing N, the player - either E or ENGO - decides to stay out from any discussion table and, hence, from any agreement, maintaining the status quo.

For both players, the utility ordering of the game outcomes that makes fully sense of this interaction is a > b > c > d. Indeed, each player can obtain the maximum payoff by exploiting the cooperation of the counterpart and taking advantage of it without shouldering any cost. Signing and carrying out an agreement when the counterparty does the same, instead, results in a lower but still good result, better than the outcomes of mutual defection. Finally, being exploited by the counterparty’s defection brings the worst payoff. Anyway, even if there is an incentive to defect from the agreement, the result of mutual defection is quite poor, and it amounts to nothing more than the ongoing status quo~~,~~ wherein each party goes on with its own strategy without considering the possibility of an agreement that may stop their ongoing conflict.

By applying Aoki's (2001) subjective games model, we assume that agents have only subjective and restricted awareness of the game's objective structure. This means that at any given time, players engage just a tiny fraction of technologically feasible actions and their combinations as "repertoires" (Dosi & Marengo, 1994 in Aoki, 2001). Indeed, the major game does include nine mini games, but only two of them are immediately visible to the participants - see Tables 1.1 and 1.2[[4]](#footnote-4). By assumption, we consider that parties have access to the same information, so that, for symmetrical information, there is an analogous change of frame. Appendix A1 depicts all the different small-framed games that could become apparent through other simplified concepts.

|  |  |  |  |
| --- | --- | --- | --- |
| **E**  **ENGO** | **S, D** | **S, C** | **N** |
| **S, D** | (c,c) | (a,d) | (c,c) |
| **S, C** | (d,a) | (b,b) | (a,d) |
| **N** | (c,c) | (d,a) | (c,c) |

**Table 1.1** The normal form of the mini-game frame as a redux of the big game in Figure 1, as it is visible to the players, according to their initial frame.

The simplified view wherein the interaction is focused on the possibility to respect or not to respect a given agreement drives agents to the idea that agreeing does not entail compliance, and hence toward a view of the game as a Prisoner Dilemma (PD). Indeed, potentially, the parties could reach a mutually advantageous deal on some terms of compromise between their preferences, such as agreeing that ENGO stops campaigning against logging industries, as this damages the latter’s reputation and causes them to lose customers, and E reduces the harvesting rate and agree to a sustainable way of managing the forest. However, at first, each player has a specific representation of the game in mind, according to which it would benefit from entering the agreement but then defecting and abusing the other’s cooperation. E has been profiting from breaching claims of compliance with specific environmental and social norms, whilst ENGO has been taking advantage from criticizing business organizations through mobilization and boycotting actions from ENGO. As a result, they both believe it is more advantageous to defect unilaterally and benefit from the other's cooperation, as in a PD, not least since they have no prior experience with fulfilled agreements and neither external restrictions nor endogenous motivations would compel them to conform. For instance, E perceives it could profit by continuing logging if ENGO anyway suspend or stop market campaigns, as the customers will be unaware of it and there would be no fallout on their reputation. ENGO, for its part, might benefit from continuing lobbying against E in the international market, even though E renounces to part of its logging activities, as this would increase the visibility of ENGO’s activity.

Even if the other party defects, it is more convenient for each player to defect too rather than to cooperate, because cooperating would entail incurring the cost of fulfilling an agreement while receiving no gain from the counterparty’s cooperation. Indeed, E would reduce logging while still suffering the consequences of ENGO’s market campaigns whereas ENGO’s would halt or discontinue campaigning without E changing their harvesting plans and tactics. Thus, according to their mental representation, the parties perceive their interaction as a PD where defecting is the only rational behavior for both players regardless of what the other player does, which means (D;D) is the only equilibrium in dominant strategies even when a more mutually advantageous agreement is achievable by playing (C;C).

Later, something occurs that reduces interpretative clarity of the game as a PD. The prevalent belief emerges that it is inconvenient to consider exploiting the counterparty's cooperation. The lobbying and involvement of international customers and investors is the trigger mechanism compelling ENGO and E to re-evaluate and update their own model of a subjective game in the GBR scenario (Affolderbach, 2011). ENGOs' efforts in foreign markets are gaining increasing awareness and influence not only among the general public but also among the corporate community, compromising E’s reputation and market ties with intermediary clients. E’s customers, that produce final goods for the final market, start considering that acquiring and transforming wood extracted in excess from forests under environmental conservation constraints might result in significant losses of reputation and reduced profit. Thus, they ask the logging companies to find reasonable solutions to the conflict, otherwise threatening the cancellation of contracts.

As a result, E finds more and more difficult to implement the defection strategy, i.e. finding intermediate customers willing to buy the excess wood at the same price, as the latter would fear to be criticized too for being complicit in irresponsible conduct. Possibly, it might find alternative buyers and markets, but this might entail selling the wood at lower prices. Thus, successful unilateral defection becomes nearly impossible, up to the point that E sees no difference between entering into the agreement and then defecting, and simply staying out of it and going on alone without trying to deceive the interlocutors. Indeed, the latter appear to be capable of draining the benefits of deceit through the effects of boycotting campaigns across the entire supply chain.

On its part, ENGO starts feeling that predatory costs for market campaigns are increasing but targeting all business companies operating in the area without differentiation based on their conduct could not bring the expected results. Indeed, intermediate customers could find themselves unable to discern the information they receive from the ENGO and see it as not reliable, thus deciding to continue buying from E. Moreover, ENGO starts seeing that E would probably accept to revise their harvesting plans only in exchange for a halt to market campaigns that damage their reputation, otherwise there will be no market effect.

In conclusion, the changing international framework bring the parties to see a different game, where each of them can either enter the agreement and try to obtain the associated benefits or, if it does not trust the counterparty, it can stay out from any negotiation table. Unilateral defection, instead, is no longer as viable and convenient a strategy as it used to be. According to their new mental models, players move their focus from a PD - Table 1.1 - to a Stag-Hunt (SH) game where the payoffs-dominant solution can prevail (Aoki, 2001; Cecchini Manara & Sacconi, 2019a) – Table 1.2. Indeed, if they do not trust each other and fear the counterpart’s defection, staying out could be a safer choice. Yet, it does not lead to an outcome as favorable as entering the negotiation and cooperating, as it implies predatory and defensive costs typical of remaining in a state of nature (Buchanan, 1975; Gauthier, 1987).

|  |  |  |  |
| --- | --- | --- | --- |
| **E**  **ENGO** | **S, D** | **S, C** | **N** |
| **S, D** | (c,c) | (a,d) | (c,c) |
| **S, C** | (d,a) | (b,b) | (d,c) |
| **N** | (c,c) | (c,d) | (c,c) |

**Table 1.2** The normal form of another mini-game frame as a redux of the big game in Figure 1, as it is visible to the players, after a shift in their initial frame.

In turn, this outcome generates a shift in the FNs-ProvGov’s subjective frames of their interaction as well, as in an institutional complementarity scenario where the institutions prevailing in a certain domain can affect choices in another domain, even though the actors in the two domains do not strategically coordinate their choices (Aoki, 2001).

Considering the game in Figure 1 and Table 1 as played between FN and ProvGov:

* By playing S-C, the player agrees to enter a G2G relation with the counterpart, i.e., to set up a decision-making table that must scrutinize the proposal by other stakeholders. The ProvGov agrees to recognize FN as another governmental actor endowed with decision-making powers and the FN agrees to work in partnership with the ProvGov.
* By playing S-D, the player agrees on the same points but defects and exploits opportunistically the other's cooperation, which for the ProvGov means undertaking decisions without involving FN while for FNs means continuing opposing and campaigning against the ProvGov, as well as threatening withdrawal of electoral support.
* By playing N, the player decides to stay out from any discussion table and, thus, from any agreement, continuing with the status quo.

As for E and ENGO, at the beginning players see just a portion of the game, the one resulting in a PD. Indeed, both FN and ProvGov perceive they could enter into an agreement and then abuse the other’s cooperative strategy. For ProvGov, this would be beneficial because it could avoid bigger compromise with industries than involving FN in decision-making would require. FN, for their part, still perceive the ProvGov as overly collusive with the logging industries, to the point where they see getting into an agreement and complying with it as embracing this collusion. Instead, they believe that unilateral defection will improve their bargaining power for an agreement that takes more account of their demands.

However, there comes a point when the rising international awareness of the issue, as well as the shift in the game between E and ENGO, alter the perception that the ProvGov and FN have of their own game. The ProvGov starts seeing the possibility to collude with logging industries and abuse the unilateral cooperation of FN as increasingly problematic, as E themselves are more careful of the effects the conflict is having on the entire supply chain worldwide, to the point that they are negotiating with ENGO. On its part, FN is concerned that its unilateral defection would push the ProvGov to negotiate the agreement with the ENGO and E, possibly with little or no regard for their claims. Thus, unilateral defection is no longer a viable strategy, since it would open the door for the other parties to reach an agreement from which FN is excluded and hence cannot benefit.

Still, both the ProvGov and FN could simply stay out of any negotiation table if they fear the counterpart’s unilateral defection; however, this would be less convenient than negotiating and entering into a G2G relation, which could instead increase their stance *vis-à-vis* ENGO and E. Indeed, both players start feeling threatened by an alliance between ENGO and E, thinking the "stakeholders table" - the one between ENGOs and FN - is taking choices on their own land (Barry, 2011; Howlett et al., 2009; Ratio, Saakiroski and Barry, 2013).

Thus, FN and ProvGov too, activate a new mental frame of their interaction, which can also be modelled as a SH game, as they do not observe anymore the (S,D ; S,C) or (S,C ; S,D) pair with asymmetric payoffs but instead see the case in which defection results in a poor outcome, which is the same as making no deal (N), as there is no benefit associated with unilateral defection but almost harm. As a result, the summary description of the game the players have in mind is reduced to one that accounts for the possibilities of either entering into the agreement and cooperating (S,C) or not entering it (N) – see Table 1.2.

**5. Compliance with the agreement: conformity preferences favoring the emergence of socially fair behavior.**

The games analyzed in Section 4 show the emergence of equilibria in the interaction between E-ENGO and FN-ProvGov, where the parties see cooperation as a feasible outcome. In the real history, this served as the basis for the big negotiating table created by the ProvGov with the aim of coming to an agreement inclusive of FN’s demands, ENGOs’ environmental concerns and E’s concerns for their reputation in the market. This exact phase could be modelled as a bargaining or deliberation game. However, it is not the focus of this paper, which instead aims at modelling the *ex-post* implementation game, i.e., the factors that could make this agreement effective even in the absence of external enforcement mechanisms.

This section deals exactly with the compliance problem, that is with *ex-post* rationality concerning adherence to the agreement (Gauthier, 1986). The implementation problem asks how an agreement can generate motivational causal forces strong enough to induce its execution in situations where at least one party could abuse it and obtain a better outcome in material terms.

In the model we present there is no complete overlapping between phase 1 - from protracted conflict to realistic prospect of feasible cooperation by agreement - and phase 2 - implementation. While in phase 1 agents play two parallel games, we model phase 2 as a four-players game between ENGO, E, international customers of E (strong stakeholders Ss) as active players and FNs as a dummy player.

In fact, the two phases represent two separate but linked problems. Market leverage by ENGO targeting Ss was essential in modifying E's framing of the game and, as a result, the perception of possible equilibria in the historical instance. As a result of their pressure on wood businesses to adjust their own strategy, Ss have indirectly played an important role in transforming a protracted conflict into viable prospects for a deal. Then, since the multi-stakeholder agreement was signed, there has been some degree of voluntary compliance. However, this is just a partial solution because many corporations have failed to share information on mapping and cutting intentions (CITE). This contradictory evidence questions the conditions under which full compliance could be reached.

Thus, we hypothesize that *ex-post* conformity with the GBR agreement could be explained by extending the psychological game developed by Degli Antoni & Sacconi (2013) in their work on Corporate Social Responsibility, so to include Ss as active players of the game. Additionally, we show how weakest parties’ organizational capacity and mobilization is decisive for a new equilibrium to emerge that is not just environmentally conscious but also socially fair. In the *ex-ante* phase, this is evident from the role that ENGO and FN play in changing the incentives of E and ProvGov respectively. In the *ex-post* phase, this is illustrated by ENGO’s capacity to boycott (CITE) in alliance with FN, i.e., to stay out and let the game end.

*5.1 Psychological Trust Game and conformity preferences: explaining ex-post compliance*

In their paper, Degli Antoni & Sacconi (2013) explain conformity through the introduction of what behavioral economists call motivational complexity. Their model considers agents that do not behave merely in their own self-interest but also have a sense of justice (Rawls, 1971) that translates into psychological utility for conformity with a principle of justice called T. When conformity preferences are incorporated, the game becomes a psychological game, and novel Psychological Nash Equilibria (PNE) might develop, i.e., scenarios in which no player has an incentive to depart from the chosen strategy when considering the counterpart's plan. Furthermore, the authors' model recognizes the importance of strong stakeholder Ss in shaping firms' incentives.

The authors start by discussing a formal model of a network (Lippert & Spagnolo, 2011) whereby an enterprise E is linked with a strong stakeholder (Ss) and two weak stakeholders (Sw1 and Sw2) who are linked one another directly or through other agents. While E and Ss have a mutual relationship and prefer to cooperate indefinitely, the relation between E and its weak stakeholders is unilateral and deficient. Sw1 and Sw2 have a low discount factor for future cooperation and thus interest in repeated cooperation, whereas E has different material incentives and thus interest in defecting. Taking this as a starting point, the authors then set up their own model to analyze the games played within the network. Specifically, they keep the relation between E and Sw as a PD, and then devise a 3-players game G (Figure 2, Degli Antoni & Sacconi, 2013:224) to model the relation between E and its stakeholders, weak and strong, where one of these is a dummy player Sw.

In G, only E and Ss are active players while Sw is a dummy one. The Ss, the ENGOs in their case, has the faculty of deciding between entering (e) or not (-e) into a cooperative relation with E, depending on whether it thinks the latter will abuse (a) or not (-a).If Ss plays e, both E and Ss can act Fairly (FE and FSs) or Unfairly (UE and USs) towards the dummy player Sw, meaning they must decide whether to collude and appropriate the entire surplus or save a part for the Sw.

Interestingly, Degli Antoni & Sacconi (2013) arrive at demonstrating that, if agents are endowed with only material preferences, (i) E and Ss will always collude and (ii) E has no incentive to cooperate with its Sw.By introducing conformity preferences instead, mutual cooperative and fair behavior can become endogenously sustainable: precisely, (i) Ss might have an incentive to boycott firms that do not cooperate with Sw and (ii) both Ss and E might have incentives to cooperate with the Sw in the long term. To understand this passage, it is necessary to assume that all parties have, at least hypothetically, participated in a pre-play communication stage in which they have agreed on a principle of justice whose implementation would entail that both E and Ss act fairly with Sw, leaving a portion of the surplus to it.

In a way, a criterion compatible with that of framed games model is at work, as the agreement may be understood as having a framing effect. By viewing the game as a situation where agreements are made, each player assumes that the other is the type of player that makes agreements and, consequently, until the proof of the contrary unravels, i.e. by default, it respects them. This presumption implies an expectation of mutual conformity that, in turn, generates a motivational factor favoring compliance, which enters the players’ utility function as a psychological component of it. Precisely, conformity preferences enter the picture when (i) an agreement is reached in a pre-play communication phase over a distributive justice principle T, such as a principle of equal treatment for the supply chains of companies like E, and (ii) this agreement activates reciprocal conformity beliefs. This last argument pertains to the fact that, in the same game, both Ss and E, by default assume that the counterpart will follow a strategy that fully realizes the distributive principle, or it is very close to doing so. Indeed, it would make no sense to agree on a principle expressing a common intentionality but not having the same intention to do it immediately, prior to any negative experience with implementation. Once these two conditions held, a disposition to conformity to the principle T is activated, which is represented in the model by a positive parameter of psychological utility 𝜆 that adds to the material payoffs of E and Ss. If this parameter is high enough to constitute a sufficient incentive for effective compliance with the agreement, compliance strategies are a new equilibrium of the game.

Ss

e

-e

Fair (FE)

Unfair (UE)

Ss

Fair (FSs)

Fair (FSs)

Unfair (USs)

Unfair (USs)

2

2

(2)

4

2

(0)

2

4

(0)

3

3

(0)

1

1

(1)

**Figure 2.** The extensive form of the game G representing the relationship between the firm and the strong stakeholder (the numbers at the edge of the three represent the payoffs obtained by Ss, E, Sw respectively) (Degli Antoni & Sacconi, 2013:221)

This model constitutes a good starting point to represent compliance with the agreement over the GBR case, for different reasons. First, it catches the essential role activism and boycotting play in shifting firms’ incentives towards adopting socially responsible behaviors (King, 2008; Zietsma & Winn, 2008). Secondly, the distinction between Ss and Sw and the introduction of a dummy player are relevant to capture the different types of relationship E has with international customers, ENGO and FN. Thirdly, we hypothesize that conformity preferences and psychological utility allows new PNE to emerge in the GBR case too, which are not only environmentally fair but also socially fair. This specification is of particular significance for the discourse on forests governance and sustainability, where false “green” solutions such as the creation of PAs are more and more criticized as new forms of racism and colonialism in the name of “environmental protection” (Mombeshora & Le Bel, 2009; Wilkie et al., 2006).

* 1. *Extending the model (the material payoffs game)*

As in the original model, the game is a kind of Trust Game (TG) where there is a firm – the logging industries E operating in the GBR – that is linked with strong and weak stakeholders. However, our model introduces an extension of the game to four players, each with its own strategies, in the following order:

* ENGO, which can decide whether to let or not the game begin by entering (e) or not (-e) the agreement, where this means entering dialogue with the other parties or continue lobbying in the market. Indeed, as in the real story, as long as ENGOs decided not to sit at the negotiating table, no agreement was reached and the conflict continued for about 20 years. In other words, ENGO can give or not trust to the other players, which means it has the option to either enter or stay out based on what it expects from E and Ss, i.e. how much wood they will extract and exchange and whether they will or will not comply with the agreement. If ENGO decides to stay out of the game and continue with the status quo, it will negatively influence the possibility for E and Ss to take advantage from their own commercial partnership. This element extends Degli Antoni and Sacconi’s (2013) model but catches the same idea of boycotting.
* International customers, i.e. business companies that buy and transform the wood and sell the final products in the market. Because they have a major interest in the running of the firm and they are indispensable for E’s surplus production, they can influence it significantly (Freeman et al., 2010). Thus, differently from Degli Antoni & Sacconi’s (2013) model where the Ss represents the ENGOs, we here consider these international business partners of E as Ss. They have two choices in the game: first, they can either buy or not buy wood from E. If they decide not to, the interaction ends at the status quo, while if they purchase some, they later have a second choice on how much to buy among three purchasing levels – corresponding to no constraints, environmental constraints only and socio-environmental constraints - also depending on the quantity of wood extracted by E - see Table 2.
* Logging companies (E), which can extract wood choosing among three quantities that respectively reflect no constraints, environmental constraints only and socio-environmental constraints – see Table 2.

Both Ss and E have an economic interest in colluding to extract and exchange large quantities of wood, but they are also subject to the effect of their accountability in the market.

* FN who appears in the game as dummy players, meaning that the game’s outcome is insensitive to its strategy and, therefore, it has no choices in it but its payoff is represented in the final vector of payments and reflects the compliance – or not – with socio-environmental constraints for the forest’s conservation.

|  |  |  |  |
| --- | --- | --- | --- |
| **Player** | **Strategy** | | |
| **nf** | **gf** | **sf** |
| **E** | extract and sell XMAX considering technological feasibilities and overall market request | extract XMED and do not extract in protected areas because of environmental policies | extract XMIN and do not extract in Conservancies because of social and environmental policies |
| **Ss** | buying as much wood as possible (XMAX) from E or any other seller that can offer the residual part to arrive at XMAX | buying less (XMED)wood because of environmental policies | buying even less (XMIN) wood because of social and environmental policies |

**Table 2.** Strategies available to E and Ss if ENGO decides to enter the game.

If ENGOs enters the game, depending on what E and Ss play, each of the four players receives a certain final payoff, the size and distribution of which is conditional on the total surplus W created through the agreement. Given the amount X\* of timber extracted and traded between E and Ss, W is a function of:

* the monetary value, or profit, P(X\*) that E and Ss obtain from legal extraction and transformation of timber into final commodities;
* the profit P#(XD) that either E or Ss derive from illegally selling on or buying from the market quantities in excess with respect to the amount allowed by the agreements of extraction limitation, i.e., the positive residue XDE = XE - XSs or XDSs = XSs - XE, where XE is the quantity extracted by E and XSs is the quantity purchased by Ss. Specifically:

0 if XSs > XE

XDE =

XE - XSs  if XE > XSs

0 if XE > XSs

XDSs =

XSs - XE  if XSs > XE

* the positive environmental externalities (U) that E, Ss and ENGO derive from an eventual reduction of = + if <, with U being a positive decreasing function of XTOT, i.e.
* the socio-environmental benefits (BSA) for FN eventually generated by E and Ss through reducing , with BSA being a positive decreasing function of XTOT.

Thus,

W = f (P(X\*), P#(XD), U, BSA)

Importantly, the profit E and Ss earn comes from the difference between revenue V and production costs C obtained by both, and it is assumed to be a positive increasing function of the extracted and exchanged quantity X\*. E obtains P(X\*) by extracting and selling a certain amount of wood X\* to Ss, while the latter derives it from the industrial transformation of wood into goods that are then sold to final consumers.

Depending on E and Ss’ strategies, there are three levels of X\*, namely XMAX > XMED > XMIN, corresponding to a non-fair, green fair and socially fair strategy and to a PMAX > PMED > PMIN respectively. Thus, profit amounts to a certain percentage of the value generated by the quantity of extracted wood and it is a linear transformation of this quantity by a parameter q. E and Ss equally share the profit they obtain from the exchange (½P). However, if E is not satisfied with what Ss buys, i.e., E extracts XMAX but Ss only buy XMIN (XE > XSs) because of socio-environmental strategies, E can still sell further wood into alternative market, possibly at a lower price. Similarly, if Ss wants to buy more wood than what E extracts (XSs > XE), it can refer to alternative logging companies that operate outside the agreement. In these cases, either E or Ss respectively, have in their profit function the added value P#(XD), as a percentage of the value of XD, generated through a linear transformation of this quantity by a parameter z. Because the economic conditions are less favorable in the secondary market, q > z.

Further, their overall utility functions contain the utility coming from environmental externalities U. To compute it, we first take the fractionary value x as an index of avoided extraction in relation to the maximum feasible one,

Such index represents the share of XMAX saved from logging and, as a function of extraction’s reduction, it increases as XTOT decreases. Then, we define the utility deriving from positive environmental externalities as an increasing but marginally decreasing function of the avoided extraction – as captured by the elevation to the factor - that becomes null for XTOT = XMAX. In a formula, we have that

The elevation to the factor assures that the payoffs of the actors are not almost entirely dependent on the utility derived from environmental externalities, but also on the one derived from private gains.

Thus, the typical utility function for E and Ss is

u*i*½P + if XD < 0, or

u*j* = ½P + P# + if XD > 0

Where

,

with q > z

and

Positive environmental externalities U also enters ENGO’s utility function. However, in this case the utility is also multiplied by a positive constant s>1: because the reduction of wood’s extraction for environmental conservation has to do with its organizational goal, ENGOs attaches to U a private payoff.

Concerning FN, instead, the utility it derives from positive environmental externalities is a function of the closest possible reduction in the level of extracted timber. Thus, maximum potential externality – weighted by a factor t – enters FN’s utility function only if the extracted timber is reduced to the minimum XMIN, due to social and environmental policies. This is justified by the fact that FN obtains real benefits – that we call socio-environmental benefits (BSE) - only when there is substantial reduction of extraction and parts of the forests are set aside from logging and converted into Conservancies, on which FN has rights of use. Thus, FN’s utility function represents the distance from the goal of minimizing XTOT, and it is calculated as

Given this, the typical utility functions for ENGO and FN are:

In substance, the subgame between E and Ss must be intended as a simultaneous move game of supply and demand of a certain quantity of extracted wood. The quantities are XMAX,  XMED and XMIN, with XMAX equals the quantity produced using full production capacity, while the last two correspond to legal constraints imposed by agreements. The parameters q, z, r, s and t transform quantities of wood into components of the players’ utility functions.

The transaction between E and Ss take place at the minimum value of X between supply and demand; yet, if E extracts more than Ss demands (XE > XSs) or *vice-versa* (XSs > XE), the residue XD might be sold into or bought from a secondary market, although at worse economic conditions. On this basis, different profit levels can be derived for the exchange in the official market (P) or the secondary market (P#). The first is a positive increasing function of the quantity X\* exchanged between E and Ss, for a factor q, while the second is a positive increasing function of the residual wood XD exchanged in the secondary market for a factor z, with q<z. Hence, a limitation in XTOT occurs only when both the parties demand or supply XMED or XMIN.

Matrix 2.1 presents a general form of the game. Where XTOT = XMAX and there is no reduction in timber extraction with respect to the maximum feasible level, and . Thus, the utility deriving from positive environmental externalities and socio-environmental benefits do not affect ENGO and FN’s payoffs, which are equal to 0. Moreover, since it is better for them to maintain the status quo than entering an agreement and being abused by E and Ss,, we conventionally set the status quo equal to 1.

-e   
ENGOs (1, 1, 1, 1)

e

|  |  |  |  |
| --- | --- | --- | --- |
| **E**  **Ss** | **nfE = XMAX** | **gfE = XMED** | **sfE = XMIN** |
| **nf = XMAX** | 0  XMAX + (XMAX –XMAX)1/2  XMAX + (XMAX –XMAX)1/2  0 | 0  XMED + z(XMAX-XMED) + (XMAX –XMAX)1/2  XMED + (XMAX –XMAX)1/2 | 0  XMIN + z(XMAX-XMIN) + (XMAX –XMAX)1/2  XMIN + (XMAX –XMAX)1/2  0 |
| **gf = XMED** | 0  XMED + (XMAX –XMAX)1/2  XMED + z(XMAX-XMED) + (XMAX –XMAX)1/2  0 | s(XMAX –XMED)1/2  XMED + (XMAX –XMED)1/2  XMED + (XMAX –XMED)1/2  t[1- | s(XMAX –XMED)1/2  XMIN + z(XMED-XMIN) + (XMAX –XMED)1/2  XMIN + (XMAX –XMED)1/2  t[1- |
| **sf = XMIN** | 0  XMIN + (XMAX –XMAX)1/2  XMIN + z(XMAX-XMIN) + (XMAX –XMAX)1/2  0 | s(XMAX –XMED)1/2  qXMIN + (XMAX –XMED)1/2  qXMIN + z(XMED-XMIN) + (XMAX –XMED)1/2  t[1- | s(XMAX –XMIN)1/2  qXMIN + (XMAX –XMIN)1/2  qXMIN + (XMAX –XMIN)1/2  t[1- |
| **no-e** | 1,1,1,1 | 1,1,1,1 | 1,1,1,1 |

**Matrix 2.1** A general example of normal form of the *ex-post* game between ENGOs, Ss and E, considering only material utility. Final payoffs are written for ENGO, Ss, Eand FN respectively.

The payoffs in the cells are ordered as follows. Considering E,

uE(e, e-nf, **nfE)** > uE(e, e-gf, **gfE**) > uE(e, e-gf, **nfE**) > uE(e, e-nf, **gfE**) = uE(e ,e-sf, **gfE**) > uE(e, e-sf, **sfE**) > uE(e, e-sf, **nfE)** > uE(e, e-gf, **sfE**) > uE(e, e-nf, **sfE**) >1 > 0.

A symmetrical order holds for Ss.

For ENGOs, instead,

uENGO(e, e-sf, sfE) > uENGO(e, e-gf, gfE) = uENGO(e, e-sf, gfE) = uENGO(e, e-gf, sfE) > 1 > 0.

The same relations hold for FNs.

The order relations between the payoffs are reasonable given the basic variables that compose them. For instance, taking (XMAX, XMAX) corresponding to (e, e-nf, nfE) and (XMED, XMED) corresponding to (e, e-gf, gfE), we have that

qXMAX ―qXMED >(XMAX –XMED)1/2

so that

qXmax qXmed + (XMAX –XMED)1/2

Meaning that the first gives E and Ss an higher payoffs than the latter because the utility deriving from positive environmental externalities does not repay the loss of profit.

In another case, instead, positive environmental externalities are high enough to counterbalance the possibility of additional profit beyond the one granted by respecting the XMED limitation of extraction. Indeed,

z(XMAX-XMED) < (XMAX –XMED)1/2

so that

qXMED + z(XMAX-XMED)< qXMED + (XMAX –XMED)1/2

meaning that (XMED, XMED) corresponding to (e, e-gf, gfE) gives E and Ss an higher utility than (XMED, XMAX) corresponding to (e, e-gf, nfE). Indeed, the former is an equilibrium outcome of the game (see below).

Analogous inequalities can explain the payoffs’ ordinal properties in each cell. For instance, taking E’s payoff for (XMIN, XMED) corresponding to (e, e-sf, gfE) and for (XMIN, XMIN) corresponding to (e, e-sf, sfE), we have that, since the extra-profit gained by selling wood in the secondary market outweighs the marginally decreasing utility deriving from increasing externality, the following condition holds:

z(XMED-XMIN) > (XMAX –XMIN)1/2 ― (XMAX –XMED)1/2.

Then,

qXMIN + z(XMED-XMIN) + (XMAX –XMED)1/2 > qXMIN + (XMAX –XMIN)1/2

meaning that E’s payoff for (XMIN, XMED), where E exceeds Ss’ demand and sells in the secondary market, is higher than its payoff for (XMIN, XMIN) corresponding to (e, e-sf, sfE). Indeed, this socio-environmental outcome is not an equilibrium of the game (see below).

gfE

nf

**Ss**

gf

sf

sf

sfE

**E**

e

no-e

no-e

e

**Ss**

**ENGO**

gf

nf

**Ss**

sf

nfE

**Ss**

nf

gf

**Figure 2.** The extensive form of the ex-post compliance game representing the relationship between ENGO, Ssand E. FN is a dummy player. The numbers represent the payoffs for ENGOs, Ss, E and FNs respectively.

Then, to consider an example of how the game is played, let’s assume XMAX = 100, XMED = 60 and XMIN = 30 and q = 0,6; z = 0,1; s = 2 and t = 0,3. It is now possible to obtain the numerical payoffs of the sub-game between E and Ss as represented in Matrix 2.1A - see Appendix A2 for complete calculations.

-e   
ENGOs (1, 1, 1, 1)

e

|  |  |  |  |
| --- | --- | --- | --- |
| **E**  **Ss** | **nfE = XMAX = 100** | **gfE = XMED = 60** | **sfE = XMIN = 30** |
| **nf = XMAX = 100** | (0, 30, 30, 0) | (0, 22, 18, 0) | (0, 16, 9, 0) |
| **gf = XMED = 60** | (0, 18, 22, 0) | (13, 24, 24, 12) | (13, 18, 15, 12) |
| **sf = XMIN = 30** | (0, 9, 16, 0) | (13, 15, 18, 12) | (17, 17, 17, 21) |
| **no-e** | 1,1,1,1 | 1,1,1,1 | 1,1,1,1 |

**Matrix 2.1A** A numerical example of normal form of the *ex-post* game between ENGOs, Ss and E, considering only material utility. Final payoffs are written for ENGOs, Ss, Eand FNs respectively.

* If both E and Ss play nf, they obtain the maximum monetary value from extraction – PMAX equal to 30 in the numerical example – and appropriate it all. No positive externality U is produced and both ENGOs and FNs are left with a payoff equal to 0.
* If both E and Ss play gf, they renounce to a part of the market so that the total monetary value they obtain – PMAX equal to 24 – is reduced and equally distributed among them. A positive environmental externality U = 6.3 is generated that enters E, Ss and E’s consequence functions, while FNs obtain an intermediate value of BSA.
* If both E and Ss play sf, they renounce to an even larger share of the market and the profit P they obtain decreases to 17 while FNs obtains the maximum BSA.

However, it is not necessarily the case that E and Ss play the same strategy and, thus, equally share the profit.

* If E plays nfE while Ss plays gf, E does not comply with the agreement and extracts XMAX, while Ss adopts an environmentally friendly strategy and renounces to part of the timber sold from E, buying only XMED. Thus, E obtains an extra profit P# equal to 4 from selling XD to an alternative market and its payoff will be higher than Ss’. Because the game is symmetric, the opposite is true when E plays gfE while Ss plays nf, as Ss obtains an additional profit P# through buying from alternative logging companies. In both cases, there will be neither U nor BSA.
* If E plays nfE while Ss plays sf, for the same logic E obtains a higher payoff than Ss, but a lower one than in the previous case, as the share of extracted timber sold in alternative market (XD) is higher, but for the same quantity, this wood gives a lower profit than the timber sold to Ss. The opposite is true if E plays sfE while Ss plays nf. In both cases, there will be neither u*i*(U) nor BSA.
* If E plays gfE while Ss plays sf, for the same reasons*,* the former makes a higher profit than the latter and *vice versa*. However, in these cases, ENGOs and FNs obtain some BSA as the amount of extracted timber is lower than in previous cases.

Given this matrix and analyzing the subgame between Ss and E, when ENGOs and Ss have already entered the game, there are two Nash equilibria.

* The first is in (no-e, e-nf, nfE) = (1,1,1,1). In the subgame between Ss and E, the equilibrium would be in (e-nf, nfE), which would bring a payoff of 30 for Ss and E but 0 for ENGOs and FNs. However, foreseeing this outcome, ENGOs would stay out of the game, so that the first equilibria is in the status quo.
* The second is in (e-gf, gfE). Indeed, if E plays gf, Ss’ best reply is to play gf and *vice-versa*. Foreseeing this, Ss would enter the game in its first move and ENGOs too. Thus, the equilibrium outcome of the overall game is in (e, e-gf, gfE) = (13, 24, 24, 12).

As mentioned above, this latter equilibrium can also be demonstrated in more general terms. Indeed, E obtains a lower payoff for (e-nf, gfE) than for (e-gf, gfE), because

qXMED + z(XMAX-XMED)< qXMED + (XMAX –XMED)1/2

Contrarily, E obtains an higher payoff for (sf, gfE) than for (sf, sfE), as

qXMIN + z(XMED-XMIN) + (XMAX –XMED)1/2 > qXMIN + (XMAX –XMIN)1/2

Because the same results hold for Ss, (e, e-gf, gfE) is an equilibrium outcome of the game while (e, e-sf, sfE) is not. This can be explained by the fact that the reduction of profits is counterbalanced by the utility derived from environmental externalities only for the environmentally fair strategy, but not for the socially fair one.

It is important to observe that (e, sf; sfE) is the outcome with the highest Nash Bargaining product but it is not a Nash equilibrium of this non cooperative subgame, since for both Ss and E a green-fair strategy is the best reply to the other player’s playing sf (sfE). Thus, the socio-environmental equitable outcome would not be sustained in equilibrium, while an equilibrium of environmental collusion between E, SS and ENGO would be feasible and it would avoid complete defection from E and SS. However, this demonstrates the possible divorce between social justice and environmental protection, which is an essential problem of sustainability (CITE). We hypothesize that the overall outcome can change by introducing conformity preferences, that allow new equilibria to emerge.

*5.3. Introducing psychological utility*

As in Degli Antoni & Sacconi (2013), the above game is only the material basis for a psychological game PG where players have conformity preferences (Grimalda & Sacconi, 2005; 2007, Sacconi & Faillo, 2010) and related first and second-order beliefs.

The first step to model the PG is to identify the T value, that is the ideal principle of fair cooperation free and rational self-interested persons would accept in a pre-play communication stage where they reason “as if” they were under a veil of ignorance (Rawls, 1971). In our case, T represents an impartial agreement over forest governance, timber extraction and surplus’ distribution. Since T is intended to represent a fair agreement, according to the relevant literature (Aoki, 2001; Binmore, 2005; Sacconi, 2000, 2006a, 2010a) T is taken to be equal to the Nash bargaining Solution (Nash, 1950;1953)

Then because of their sense of justice (Rawls, 1971), players can obtain psychological utility by mutual compliance with T; this utility adds to the material one and modifies players’ final payoffs. However, the incentive to conform is subject to players’ mutual beliefs about reciprocal conformity; depending on this, new PNEs might emerge.

Thus, the second step is to include ideal utility and reciprocal beliefs in the agents’ utility functions following the psychological game concept (Geanakoplos et al., 1989; Grimalda & Sacconi, 2005, 2007; Rabin, 1993). Each overall utility function is computed as

where (𝜎) represents the material utility player *i* attaches to the state of affair 𝜎 while F[T(σ)] represents *i*’’s conformity reason given her expectations of reciprocal conformity. is an exogenous motivational parameter that measures a player *i*’s desire to conform with T while the function F[T(*σ*)] captures beliefs’ role in affecting ideal utility. Thus, given a certain state of affair *σ*, a system of beliefs b which includes player *i*’’s conditional conformity index 𝑓𝑖 and a reciprocal expected conformity index 𝑓*j*, F[T(*σ*)] measures an agent’s conformity to the agreed principle T, given what she expects the other player will play and what she expects the other expects from her. As explained more into details in Appendix A3, this can also be expressed as

𝑉𝑖 (𝜎) = 𝑈𝑖 (𝜎) + Ψ𝑖 (𝜎, 𝑇, 𝜆, 𝑏)

or, more explicitly, as

𝑉𝑖 (𝜎𝑖, b𝑖1, b𝑖2) = 𝑈𝑖 (𝜎𝑖, b𝑖1) + λ𝑖[1 + 𝑓̃*j*(b𝑖1, b𝑖2)][1+ *fi*(𝜎𝑖, b𝑖1)].

5.4 *The new psychological game*

Given the theory, we apply it to the case at point by revisiting the game we just described, assuming that ENGOs, E and Ss are endowed with conformity preferences after having agreed on a fair agreement. To do so and identify the agreement to which they will subscribe, we first need to compute T values for each state of affairs, which define the ordering of outcomes based on their compliance with the principle of justice that parties have agreed to (Grimalda & Sacconi, 2005, 2007). Precisely, we compute them for each state of affair belonging to the subgame wherein only Ss and E participate, assuming ENGO have first chosen to enter.

|  |  |
| --- | --- |
| 𝜎 | T |
| (e, nf; nfE)  (e, gf; gfE)  (e, sf; sfE)  (e, nf; gfE)  (e, gf; nfE)  (e, nf; sfE)  (e, sf; nfE)  (e, sf; gfE)  (e, gf; sfE)  (-e; nfE)  (-e; gfE)  (-e; sfE) | 0  89856  103173  0  0  0  0  42120  42120  1  1  1 |

**Table 2.2.** T values for each state of affairs of the subgame between Ss and E.

As Table 2.2 shows, for each strategy played by Ss, T assumes the maximum value for (e, sf;sfE) = 103173, intermediate values for (e, gf; gfE) = 89856, (e, gf; sfE) or (e, sf; gfE) = 42120 and (no-e) = 1, and minimum value T = 0 for all the other strategies’ combinations. This means that, under an ex-ante agreement, parties would choose (e, sf;sfE), which is the ideal state of affairs (Grimalda & Sacconi, 2005, 2007).

Then, we compute Ss and E’s overall conformity indexes to the principle T, given their expectations on the other player’s actions. For instance, if Ss foresees that E plays sfE, then Ss’ strategy to maximize T is e-sf. The same applies reciprocally: if E foresees that Ss plays e-sf, E’s strategy to maximize T will be sfE. Alternatively, if Ss foresees that E plays gfE, then Ss’ strategy to maximize T is e-gf. The same applies reciprocally in this case too: if E foresees that Ss plays e-gf, E’s strategy to maximize T will be gfE. Table 2.2 summarizes the results and Appendix A3 contains details for the entire calculation.

|  |  |  |
| --- | --- | --- |
| 𝜎 | (1 + *f*Ss) (1 + 𝑓̃E) | F(T) |
| (e, nf; nfE)  (e, gf; gfE)  (e, sf; sfE)  (e, nf; gfE)  (e, gf; nfE)  (e, nf; sfE)  (e, sf; nfE)  (e, sf; gfE)  (e, gf; sfE)  (-e; nfE)  (-e; gfE)  (-e; sfE) | (1 - 1) (1 + 0)  (1 + 0) (1 + 0)  (1 + 0) (1 + 0)  (1 – 1) (1 + 0)  (1 – 1) (1 – 1)  (1 – 1) (1+ 0)  (1 – 1) (1 – 1)  (1 – 0.6) (1 – 0.5)  (1 – 0.5) (1 – 0.6)  (1 + 0) (1 + 0)  (1 – 0,99) (1 + 0)  (1 – 0,99) (1 + 0) | 0  1  1  0  0  0  0  0,2  0.2  1  0.01  0.01 |

**Table 2.3.** Ss and E’s overall conditional conformity indexes with the principle of justice T.

By now, we can compute the overall utility function for each state of affairs, as represented in Matrix 2.4.

|  |  |  |  |
| --- | --- | --- | --- |
| **E**  **Ss** | **nfE = XMAX = 100** | **gfE = XMED = 60** | **sfE = XMIN = 30** |
| **nf = XMAX = 100** | (0, 30, 30, 0) | (0, 22, 18, 0) | (0, 16, 9, 0) |
| **gf = XMED = 60** | (0, 18, 22, 0) | (13, 24+λss, 24+λE, 12) | (13, 18+0.2λss, 15+0.2λE, 12) |
| **sf = XMIN = 30** | (0, 9, 16, 0) | (13, 15+0.2λss, 18++0.2λE, 12) | (17, 17+λss, 17+λE, 21) |
| **no-e** | (1, 1+λss,1+ λE,1) | (1, 1+0.01λss, 1+0.01λE, 1) | (1, 1+0.01λss, 1+0.01λE, 1) |

**Matrix 2.4.** The game in normal form representing the *ex-post* game between ENGOs, Ss and E, considering also psychological utilities. Final payoffs are written for ENGOs, Ss, Eand FNs respectively.

As matrix 2.4 shows, E’s overall utility functions for each state of affair 𝜎 of the subgame and hence the corresponding payoffs, are symmetrical to player Ss’ ones. What emerges clearly from this analysis, albeit partial, is that depending on the absolute value λ assumes, the calculus of players Ss and E’s best responses may change with respect to the game where only material payoffs were considered.

In the subgame between E and Ss, their choices are based on mutual conformity expectations, whereas they have no expectations of ENGO, since its choice is assumed. Hence, there is no need to compute the same indices in relation to ENGO's decision not to enter. Contrarily, when considering ENGO's choice, ENGO has expectations on Ss and E's further moves and on their level of conformity to the agreement that maximizes T. Thus, it is important to analyze ENGO's indices for its two options (e) and (no-e) in relation to every possible combination of Ss&E's strategies.

Accordingly, we now consider all the possible states of affair *σ* of the overall game. At any rate, to favor the reader, we will continue to present them as cells of table in which columns entries represent ENGO’s alternative choices to enter or not to enter while rows entries display each possible combination of Ss&E’s strategies considered as conjoint events, as indicated by the symbol of conjunction “&”.

Table 2.5 shows:

* T values for each cell of the game which results from the combination of ENGO’s choice with the combined action by Ss&E;
* indexes of expected and joint conformity (1 + 𝑓̃Ss,E) of players Ss&E as estimated by ENGOs considering the hypothesis that Ss&E use each of their strategies combination under the expectation that ENGOs uses each of its two strategies alternatively;
* conformity indexes (1 + *fENGOs*) conditional on the hypothesis that Ss&E use each of their strategies’ combinations considered as conjoint events;
* composite conformity indexes - conditional and expected - with respect to each possible pair of predictions.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **ENGO**  **Ss &E** | e | no-e | **e**  (1 + 𝑓̃Ss,E) (1 + *fENGOs*) F(T) | **no-e**  (1 + 𝑓̃Ss,E) (1 + *fENGOs*) F(T) |
| (e,nf & nfE) | T=0 | T=1 | (1 – 1) (1 - 1) 0 | (1 + 0) (1 + 0) 1 |
| (e,gf & gfE ) | T=89856 | T=1 | (1 – 0.12) (1 + 0) 0,88 | (1 + 0) (1 - 1) 0 |
| (e,sf & sfE) | T=103173 | T=1 | (1 + 0) (1 + 0) 1 | (1 + 0) (1 - 1) 0 |
| (e,nf & gfE) | T=0 | T=1 | (1 – 1) (1 – 1) 0 | (1 + 0) (1 + 0) 1 |
| (e,gf & nfE) | T=0 | T=1 | (1 – 1) (1 – 1) 0 | (1 + 0) (1 + 0) 1 |
| (e,nf & sfE) | T=0 | T=1 | (1 – 1) (1 – 1) 0 | (1 + 0) (1 - 1) 0 |
| (e,sf & nfE) | T=0 | T=1 | (1 – 1) (1 – 1) 0 | (1 + 0) (1 - 1) 0 |
| (e,gf & sfE) | T=42120 | T=1 | (1 – 0.6) (1 + 0) 0.4 | (1 + 0) (1 - 1) 0 |
| (e,sf & gfE) | T=42120 | T=1 | (1 – 0.6) (1 + 0) 0.4 | (1 + 0) (1 - 1) 0 |
| (no-e & nfE) | T=1 | T=1 | (1 – 0. 99) (1 + 0) 0.01 | (1 + 0) (1 + 0) 1 |
| (no-e & gfE) | T=1 | T=1 | (1 – 0. 99) (1 + 0) 0.01 | (1 + 0) (1 + 0) 1 |
| (no-e & sfE) | T=1 | T=1 | (1 – 0. 99) (1 + 0) 0.01 | (1 + 0) (1 + 0) 1 |

**Table 2.5.** Overall conformity indexes of ENGOs, for each of the two strategies at its disposal, conditional on the hypothesis that Ss&E use each of their strategies’ combinations considered as conjoint events.

Then, Matrix 2.6 displays the psychological payoffs that ENGO’s obtain for the choices to either enter or stay out, corresponding to each state of the subgame in which Ss&E use each combination of strategies. ENGO's psychological utility depends on its own conditional conformity to the prediction of a certain combination of Ss&E’s choices, as well as on the conformity of Ss&E according to their prediction of ENGO's choice. For instance, only when ENGO expects that Ss&E will play (e,sf & sfE), then the full value of λENGO enters into ENGO’s payoffs if it decides to enter, while only a partial value of it enters when Ss&E play (e,gf & gfE). For all other strategies of Ss&E, no-e gives the highest value of λENGOs, as keeping the status quo is preferable than entering and being abused.

These results reflect the idea of boycotting and strategic activism (Baron & Diermeir, 2007; Degli Antoni & Sacconi, 2015; Zietsma, 2008).Assuming ENGOs are driven by environmental concerns and care about FN, meaning their payoff function includes conformity preferences, they reward enterprises who accept the demand to revise their harvesting activities and follow through on this pledge, while they punish firms that do not (Baron & Diermeir, 2007; Degli Antoni & Sacconi, 2015). In the case at issue, rewarding means entering the agreement and comply to halt market campaigns against those industries that comply with the socio-environmental fair strategy. Punishing instead means carrying on campaigns against unfairly playing industries, and increasing their production costs by raising defensive costs (Buchanan & Tullock, 1962).

|  |  |  |
| --- | --- | --- |
| **ENGO**  **Ss &E** | **e** | **no-e** |
| (e,nf & nfE) | 0 | 1 + λENGOs |
| (e,gf & gfE) | 13 + 0,88λENGOs | 1 |
| (e,sf & sfE) | 17 + λENGOs | 1 |
| (e,nf & gfE) | 0 | 1 + λENGOs |
| (e,gf & nfE) | 0 | 1 + λENGOs |
| (e,nf & sfE) | 0 | 1 |
| (e,sf & nfE) | 0 | 1 |
| (e,gf & sfE) | 13 + 0.4λENGOs | 1 |
| (e,sf & gfE) | 13 + 0.4λENGOs | 1 |
| (no-e & nfE) | 1 + 0.01λENGOs | 1 + λENGOs |
| (no-e & gfE) | 1 + 0.01λENGOs | 1 + λENGOs |
| (no-e & sfE) | 1 + 0.01λENGOs | 1 + λENGOs |

**Matrix 2.6.** Psychological payoffs of ENGOfor the choices to either enter or stay out, corresponding to each state of the subgame in which Ss&E use each combination of strategies.

Finally, we consider psychological payoffs that the three players obtain under the hypothesis that ENGO does not enter. Clearly, these are the same for each player, given that their material payoffs (1,1,1,(1)) are identical as is their degree of conformity resulting from ENGO’s decision not to enter. In Matrix 2.6, these payoffs are represented in the upper matrix. The lower matrix instead represents the psychological payoffs when ENGO decides to enter, and Ss&E’s combination of strategies is therefore effective. Overall, Fig. 2.6 represents the psychological payoffs of the new psychological game PG’’ and illustrates strategies’ combinations of all players, for which the psychological parameter λ enters their payoffs.

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Ss & E**  **ENGO** | **e,nf & nfE** | **e,nf & gfE** | **e,nf &**  **sfE** | **e,gf & nfE** | **e,gf &**  **gfE** | **e,gf & sfE** | **e,sf & nfE** | **e,sf & gfE** | **e,sf & sfE** | **no-e** |
| **no-e** | 1 + λENGOs  1 + λSs  1 + λE  (1) | 1 + λENGOs  1 + λSs  1 + λE  (1) | 1  1  1  (1) | 1 + λENGOs  1 + λSs  1 + λE  (1) | 1  1 + λSs  1 + λE  (1) | 1  1  1  (1) | 1  1  1  (1) | 1  1  1  (1) | 1  1  1  (1) | 1 + λENGOs  1 + λSs  1 + λE  (1) |

**e**

**no-e**

**ENGO**

|  |  |  |  |
| --- | --- | --- | --- |
| **E**  **Ss** | **nfE = XMAX= 100** | **gfE = XMED = 60** | **sfE = XMIN = 30** |
| **nf = XMAX = 100** | 0, 30, 30, (0) | 0, 22, 18, (0) | (0, 16, 9, (0) |
| **gf = XMED = 60** | 0, 18, 22, (0) | 13+0,88λENGOS, 24+λSs, 24+λE, (12) | 13+0,4λENGOS, 18+0,2λSs, 15+0,2λE, (12) |
| **sf = XMIN = 30** | (0, 9, 16, 0) | 13+0,4λENGOs, 15+0,2λSs, 18+0,2λE,(12) | 17+λENGOS, 17+λSs, 17+λE, (21) |
| **no-e** | 1+0.01λENGOs,1+0.01λSs,1+0.01λE,(1) | 1+0.01λENGOs,1+0.01λSs,1+0.01λE, (1) | 1+0.01λENGOs,1+0.01λSs,1+0.01λE,(1) |

**Matrix 2.7.** The normal form of the game PG’’ representing the relationship between ENGOs, Ss and E, considering also psychological utilities.

The new PG’’ in Fig 2.7 shows that:

* If E plays sfE, Ss’ best response is (e-sf), and
* If ENGO believes that if it enters, Ss and E will then either i) collude on a non-fair or environmentally fair strategies combination or ii) play either gf and nfE or *vice versa*, its best response would be to stay out of the game. This result is particularly interesting because it completely overturns the outcome of the game where only material utilities were considered: in that game, indeed, given the same expectation ENGOs enter the game and (e; e-gf; gfE) is a weak Nash equilibrium of the game. It is therefore evident here how an agreement can activate reasons to start a relation whereby a surplus is created and shared in a fairer way.

In line with the findings of Degli Antoni & Sacconi (2013:223), these results show “a very important consequence of the conformist preference model: staying out of an unfair cooperative relation can induce the relative best level of conformity if the “cooperative” choice is such that acceding to such a proposal of unfair behavior (collusion) induces a lower level of principle T achievement (e.g. a lower Nash Bargaining Product value)”.

*2.3 New Psychological Equilibria*

To summarize, new Psychological Nash Equilibria emerge when λ enters players’ psychological payoffs. One is clearly sub-optimal (–e, nf, e-nf) = (1+λENGOs, 1+λSs, 1+λE, (1+λENGOs, 1+λSs, 1+λE, (1)) but it is nevertheless interesting because it psychologically rewards boycotting and sanctioning over green collusion. Indeed, ENGOs obtain higher psychological utility from staying out if it predicts that by entering, Ss&E would then decide for a strategies’ combination that determine low T values. Contrarily, if ENGOs predicts that Ss and E will determine high T values, its λ increases by playing e.

The other two equilibria, which correspond to the ENGOs’ entry strategy and Ss&E environmentally and socially fair strategies, are mutually non-Pareto dominant, and they generate a genuine multiplicity problem, that is a mixed-interest coordination problem. However, the socially fair outcome is more equitable in terms of payoff symmetry and Nash equilibrium. Thus, in the psychological game PG’’ this is an *ex-post* equilibrium if players’ motivational weight 𝜆 is high enough to counterbalance their temptation to adopt a non-fair or just environmentally fair strategy.

Let’s consider the outcomes of the subgame between Ss and E, assuming ENGO enters, in a descending order based on T, as shown in Table 2.8.

|  |  |
| --- | --- |
| 𝜎 | T |
| (e, sf; sfE) | 103173 |
| (e, gf; gfE) | 89856 |
| (e, sf; gfE) | 42120 |
| (e, gf; sfE) | 42120 |
| (-e; nfE) | 1 |
| (-e; gfE) | 1 |
| (-e; sfE) | 1 |
| (e, nf; gfE) | 0 |
| (e, gf; nfE) | 0 |
| (e, nf; sfE) | 0 |
| (e, sf; nfE) | 0 |

**Table 2.8.** Payoffs of the subgame between Ss and E presented in a descending order based on T

Then, the psychological utility through factor λ enters the payoffs only in certain cases. Namely,

* there is no psychological utility – and therefore no λ - when ENGO and FN obtain a payoff = 0;
* λ = 1 in a player’s payoff only for one outcome given the opponent’s strategy. This is because there is only one row associated with the maximum value of the conformity index for each column;
* λ = 1 for the outcome where T = TMAX in the matrix.

Thus, even though

uSs (e-gf, gfE) > uSs (e-sf, sfE)

and uE (e-gf, gfE) > uE (e-sf, sfE)

We can have that

VSs (e-sf, sfE) > VSs (e-gf, sfE)

and VE (e-sf, sfE) > VE (e-gf, sfE)

if and only if

λSs― xλSs > uSs (e-gf, sfE) - uSs (e-sf, sfE)

and λE― xλE > uE (e-gfE, sf) – uE (e-sfE, gf)

Where x is the weight of λ.

If these conditions hold and given appropriate players’ first and second order beliefs in mutual conformity with the ideal principle T, then playing e-sf (sfE) is the best strategy to the other playing sfE(e-sf) and the socially-fair outcome is an equilibrium of the psychological game.

In the numerical example, this means that

λSs - 0,2λSs > 18 – 17

0,8λSs > 1 ⬄ λSs > 1,25

and

λE - 0,2λE > 18 – 17

0,8λE > 1 ⬄ λSs > 1,25

Thus, when λSs > 1,25; ENGOs believe that Ss and E play (e-sf; sfE), Ss believes that E plays sfE and E believes that Ss plays (e;sf) and; each of them has a second (and higher) order belief that the other has exactly these beliefs, then (e, sf; sfE) becomes a PNE and the agreement is sustained in equilibrium. This is equal to say that ENGOs prefers not to boycott and Ss and E prefer not to collude and ensure a fair division of the surplus that also goes at the benefit of FNs.

1. https://www2.gov.bc.ca/gov/content/environment/natural-resource-stewardship/great-bear-rainforest/gbr-agreement-highlights [↑](#footnote-ref-1)
2. A more advanced technique may possibly simulate all these stages as part of a single game made up of multiple stages. However, at the time, this option appears to involve an unreasonable level of intricacy for the specified participants as well as the theoretician. [↑](#footnote-ref-2)
3. In the real case, certain environmental organizations such as Greenpeace, ForestEthics, the Sierra Club, Rainforest Action Network and Natural Resources Defense Council carried forward a joint market campaign (Armstrong, 2009) and joined together the JSP. Similarly, various business organizations spoke as a unique voice both when they initially defended their practices and their contribution to the regional economic developed and when they later decided to shift their strategies to sit at the table with ENGOs (Armstrong, 2009). FNs too formed formal coalitions, such as the Coastal First Nations Turning Point Initiative (CFN), the Skeena First Nations Stewardship Council and the Nanwakolas Council (Raitio & Saakiroski, 2014). [↑](#footnote-ref-3)
4. For a more complete presentation of the idea of mini games as alternative game frames see Cecchini Manara Sacconi, 2019a to which this part is indebted. [↑](#footnote-ref-4)