UNIVERSIDAD DE CHILE
FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS
DEPARTAMENTO DE INGENIERÍA INDUSTRIAL

NORMAS SOCIALES, COOPERACIÓN Y LIDERAZGO

TESIS PARA OPTAR AL GRADO DE MAGÍSTER EN ECONOMÍA APLICADA

FELIPE BERSALO CARRERA GALLEGUILLOS

PROFESOR GUÍA:
JUAN ESCOBAR CASTRO

MIEMBROS DE LA COMISIÓN:
ELTON DUSHA
NICOLÁS FIGUEROA GONZÁLEZ

SANTIAGO DE CHILE
MAYO 2014

Resumen

En este trabajo son estudiados los problemas de incentivos que enfrentan líderes dentro de una comunidad y sus efectos sobre la cooperación que se puede sostener en equilibrio.

Es considerado un modelo de juego repetido en el que una comunidad formada por grupos disjuntos presenta una estructura simple de liderazgo. Cada agente del juego es emparejado en cada etapa con otro jugador de cualquier grupo mediante una función de emparejamiento uniforme, tras lo cual ambos agentes juegan un Dilema del Prisionero. Se ha propuesto en este caso una norma social de dos fases, similar a la planteada por Ellison (1994), la que es implementada a través de las instrucciones entregadas por los líderes de cada grupo a sus constituyentes. A su vez, los líderes resuelven el problema de agencia que se genera entre ellos usando una estrategia de revisión simétrica basada en Radner (1985).

La literatura en normas sociales ha estudiado los problemas de incentivos que enfrentan agentes pertenecientes a diferentes grupos al interactuar entre sí en contextos fuera del marco legal (Fearon & Laitin, 1996; Greif, 1993). Además, ha identificado en algunos casos el papel que cumplen estructuras de liderazgo en la implementación de normas sociales que regulan interacciones económicas inter-grupales (Milgrom, North, et al, 1990; Fearon & Laitin, 1996; Geertz, 1973). Sin embargo, dentro de la literatura no había sido modelado explícitamente el problema de incentivos que líderes de distintos grupos enfrentan al implementar una norma social común.

Es definido un concepto de equilibrio en estrategias simétricas para un modelo de interacción inter-grupal con líderes locales. Entre las contribuciones del modelo están el mostrar cómo una simple estructura de liderazgo puede sustentar un equilibrio en el cual el comportamiento oportunista por parte de los líderes se mantiene bajo control y las interacciones inter-grupales permanecen cooperativas en el camino del equilibrio.

Un resultado significativo es que el equilibrio simétrico para los líderes es robusto. Esto significa que después de un desvío, en un número finito de etapas, es restaurado un perfil de acciones en el que todos los agentes de la comunidad cooperan entre sí. Este resultado permite darle racionalidad a la existencia de líderes, ya que estos permitirían implementar una norma social que le genera a cada agente una utilidad arbitrariamente cercana a la óptima social mientras, al mismo tiempo, se obtiene cierto grado de estabilidad pues la cooperación puede resistir desviaciones ocasionales o errores.

A las 2 Mónicas.

Agradecimientos

Quiero agradecer sinceramente a todos aquellos que han colaborado conmigo, de una u otra forma, en la consecución de mi magíster en economía aplicada en general, y en particular, en la realización de este proyecto de título. Tengo una deuda muy grande con todos ustedes.

Debo partir recordando a mi familia, especialmente a mi madre y a Mónica, mi hermana, por apoyarme desde lo cotidiano durante todo este tiempo. A Kate, mi novia, por su incondicionalidad y ganas de escuchar, con mucha paciencia, acerca de los avances y retrocesos de este proyecto y por siempre estar ahí dandome ánimo y alegría; y a mi hermano Patricio y a mi padre.

Tengo en especial una deuda muy grande con mi profesor guía, Juan Escobar, por mostrarme desde el inicio la forma correcta de abordar un proyecto de investigación; por su capacidad para orientarme y desafiarme, dándome al mismo tiempo la autonomía necesaria para poder obtener el máximo aprendizaje posible de esta tesis; por todos sus consejos, dentro y fuera del contexto de este proyecto, y en general, por su paciencia y generosidad. Debo agradecer también a los otros miembros de la comisión, Nicolás Figueroa y Elton Dusha, por su buena disposición, tiempo y valiosos comentarios; y al resto de mis profesores del magcea, en particular a Alejandra Mizala, Benjamín Villena y Leonardo Basso, por todo el apoyo, tiempo y sabiduría que me han brindado durante este tiempo.

Para terminar, tengo que mencionar y agradecer de forma especial a mis compañeros del magcea, sin los cuales haber hecho esta tesis hubiera sido un proceso mucho más difícil y menos gratificante. En particular tengo que agradecer por su entusiasmo y confianza a Nicolás Inostroza, Carlos Lizama, Diana MacDonald, Felipe Díaz, Oscar Arias, Pablo Cuellar y Stéfano Banfi.

Table of Contents

List of Figures

Introduction

0.1. Motivation

This paper studies leadership in the context of a self-governance model. The central motivation of this work is to study the incentive problem of group leaders towards each other, in a community in which disjoint groups exist. The latter causes an asymmetry of information between agents that belong to a same group and agents that belong to different groups, reducing the likelihood of conducting transactions if the probability of facing the same agent again is low. This setting resembles many real life situations since the fact that the groups are disjoint might be caused by religious, linguistic, or ethnic differences.

The economic literature on self-governance and social norms has progressed in understanding the challenges a community faces to implement economic cooperation in the absence of the rule of state law. In many of the applications of social norms to real situations, one distinctive feature is the existence of different "types" or groups defined either by ethnic, religious, linguistic, or geographic boundaries, generating asymmetric information between the agents. Moreover, when discussing how groups of agents can sustain inter-group cooperation in the absence of law, casual observation points out that group leaders or mediators, for example religious or community leaders, can play an important role in the implementation of social norms (Geezer, 1973).

That empirical regularity introduces interesting questions as to why a human community would find it rational or desirable to have their political or religious leaders central in the execution of its economic relations with the rest of the world; however, the incentive problems faced by group leaders or mediators in inter-group relations have not yet been addressed by the literature. Important asymmetries of information may arise in the interaction between group leaders, opening the door for opportunistic behavior on the part of one group leader against the rest which could make it difficult to achieve cooperation on equilibrium.

In the present work, we study a repeated game model in which several disjoint groups of agents interact using a common social norm. Every group has a distinct agent that acts as group leader who, to implement the social norm and depending on the history of the game, has to give different instructions to the members of his group allowing for the possibility of opportunistic behavior on his part. In order to prevent this, each group leader will continuously supervise the instructions the other group leaders are giving.

The proposed model can be though as having two sections. The first one would correspond to the repeated interactions between the non-leader agents. These are regulated by the social norm whose implementation uses the group leaders. This part of the model uses a two-phase social norm adapting the work of Ellison (1994); although here the model considers a continuum of agents and each group of agents has an independent random device to regulate the transition between phases. In this section of the game, we have to make sure that for the non-leader agents it is incentive-compatible to participate in a long term economic relation with the other groups that is regulated by the proposed social norm.

The second section of the game, on the other hand, is the problem of the group leaders. Since the transition between phases in the social norm is probabilistic, and under the assumption that the group leaders payoffs are proportional to the average stage payoffs their group members obtain, group leaders can have opportunities for unilateral profitable deviations from giving fully honest instructions. This configures a symmetric moral hazard problem in which we can characterize each group leader as an agent of the rest of the group leaders. Clearly this is a symmetric problem; therefore, each group leader is simultaneously agent and principal of every other group leader in the model.

To solve the problem of the group leaders we rely upon the notion that by the law of the great numbers in the long run the behavior of the group leaders should converge to the expected behavior of a fully honest group leader. Thus, the solution of the group leaders' moral hazard problem is made finding the parameters of a symmetric review strategy, defined using the methodology developed by Radner (1985), and showing its optimality against an optimal deviation of the group leaders acting as an agent of the rest.

Among the contributions of this work are to explicitly model the incentive problem faced by leaders in inter-group relations in a self-government model, and to define a suitable concept of equilibrium under symmetric strategies for the problem that shows how a simple structure of leadership can successfully implement a social norm in which cooperation can be sustained despite the above mentioned moral hazard problem. Moreover this symmetric equilibrium is robust, meaning that it takes, in expected value, a finite number of stages until the full cooperation is restored. This result justifies the existence of community leaders in a multi-group setting since it can be argued that their presence makes it possible to implement social norms that generate a higher degree of stability.

This paper is organize as follows. Chapter I describes the model and the main assumptions. A simplified version of the model, in which group leaders are left momentarily out of the picture is presented in Chapter II in order to study the incentive compatibility conditions for the non-leader agents and to establish a benchmark for the general model. In Chapter III, the main results regarding the model with group leaders are presented. Finally, we present some important conclusions and possible futures lines of research.

0.2. Related literature

There is a growing literature on social norms and inter-group cooperation, including papers not only from Economics but also from other social sciences as Sociology and Anthropology.

In particular, the above mentioned economic literature starts with the seminal papers of Okuno-Fujiwara & Postlewaite (1995), and Kandori (1992). The latter studies a model of repeated game with random matching and shows that an efficient equilibrium, one in which cooperation in a prisoner's dilemma is always observed on-the-path, exists for two extreme information structures: one case in which there is no transmission of information at all, but the strategy followed by the players generates a contagion effect in case a deviation occurs; and a second one, in which there is an agency that has perfect information about every history and assigns labels honestly to all the players of the game. The properties of models of repeated games with random matching are further studied in the model of random matching by Elisson (1994). In this paper, he shows the existence of equilibrium in a model with anonymous random matching, in which the strategy followed by the agents considers a "punishment phase" and the duration of the punishments is determined using a public random device.

Gosh & Ray (1996) presents a repeated game model without information flows in which agents are divided in two types depending on their discount factor and can choose to keep having interactions with the same "partner" in the futures stages of the game. Finally, Acemoglu & Jackson (2012) studies the dynamics of social norms in a repeated game model with overlapping generations.

Greif (1993) focuses on the Maghribi traders in the Mediterranean basin, explaining how an ethnic community, distributed along very distant cities, could establish and sustain a successful commercial web despite the existence of options for opportunistic behaviour given its geographic dispersion.

In the case of groups belonging to different ethnic backgrounds, Fearon & Laitin (1996) present two possible strategies that interacting ethnic groups can use to maintain cooperation between their members, both of which require the capability of coordinating punishments directed at deviating players belonging to their own group, or towards the other ethnic group. Some case studies presented in the same article stress the importance of group leaders for the implementation of these strategies. For example, in the Ottoman Empire, religious and community leaders of the main minorities (Greek Christians, Jews, and Armenians) received a high degree of autonomy as long as they were effective in monitoring the members of their communities in their interactions with outsiders and dealing with deviating agents.

In other cases, mediators help regulate inter-group relations. For instance, Geezer (1973), also cited by Fearon & Laitin, explains how the trading relations between Jews and Berbers in Morocco were helped by the presence of Berber mediators that used their informal power within their group to act as "insurers", having the capacity to compensate for damages caused by deviating members of that community.

Chapter 1

The Model

1.1. Main elements

The model studied is one of repeated games with contagion. A community is formed by a continuum of agents, where each agent is indexed by i, such that $\mathrm{i} \in [0, 1]$, and this continuum of agents is arranged as a circumference.

	C	D
C	$1, 1$	$-l, 1+g$
D	$1+g, -l$	$0, 0$

Figure 1.1: Matrix of payments of the stage game

There exists a uniform matching function that distributes, at the beginning of each stage, all the agents in pairs. Formally, the agent assigned to the player indexed by i in the stage t of the game is denoted by $m_{\mathrm{i}}(t)$. The value of $m_{\mathrm{i}}(t)$ is known to every player i, that is, every player knows the identity of the agent that was assigned to him in every stage $t \in \{0, \ldots, \infty\}$ after the realization of the match. After this assignation, each pair of agents play a prisoner's dilemma with the usual structure of payments (g, $l > 0$), as shown in Figure (1,1).

Every agent in the game belongs to one of N disjoint groups or segments of players (see Figure 1.2). Each group is endowed with a very simple structure of leadership that can be associated with an ethnic, political or religious leadership. This structure of leadership will consist of a single agent per group (who will be called in what follows the *group leader*) that does not participate in the above described random matching. Finally, every agent in the game (including group leaders) discount future payments by discount factor δ and their utility corresponds to the discounted sum of the payments they receive during the game.

We will make the assumption that the stage payment a group leader receives in any stage

t of the game is simply a lineal function of the average stage payments of the members of his group of players: $u_L^k = a + b \cdot P_t^k$. Where k is the index of the group, P_t^k is the average payment of the members of group k in stage t, a is a minimum payment (potentially zero) and b is a proportionality constant $(a, b > 0)$.

The group leader receives, at the end of every stage t of the game, information from every player that belongs to his group about the identity of the agent he was matched with and the actions played by both of them in that stage. We assume the information the non-leader players transmit is truthful. Moreover, the group leader is the only player in the group that can see the realizations of an independent random device available for each group.

Finally, at the end of each stage (and after he has observed the realization of the random device), the group leader sends the same signal s_t^k to every member of his group, telling them how to play in the stage $(t+1)$ of the game and giving them the list of agents of group k that have deviated until stage t (this component of S_t^k will be explained below).

$$S_t^k = D_t^k \cup A_t^k$$

where $D_t^k \in \mathbb{R}^n$ and $A_t^k \in \mathbb{R}^{(N-1)}$

The vector A_t^k is a vector of "instructions" with $(N-1)$ components, one per each of the other groups.

$$A_t^k = \left\{ A_t^{k,1}, A_t^{k,2}, \ldots, A_t^{k,k-1}, A_t^{k,k+1}, \ldots, A_t^{k,N} \right\}$$

Where $A_t^{k,l}$ is the action group leader of group k tell his members to do in stage $(t+1)$, regarding members of group l (where $k \neq l$, and $k,\ l \in \{1, \ldots, N\}$). We will refer to the vector A_t^k below as *communication action.*

Lastly, notice that $A_t^{k,l} \in \{C, D\}$, $\forall k,\ l \in \{1, \ldots, N\}$. For instance, if $A_t^{k,l} = C$ then when faced by a member of group l in stage $(t+1)$ and agent belonging to k is told to play the stage action C.

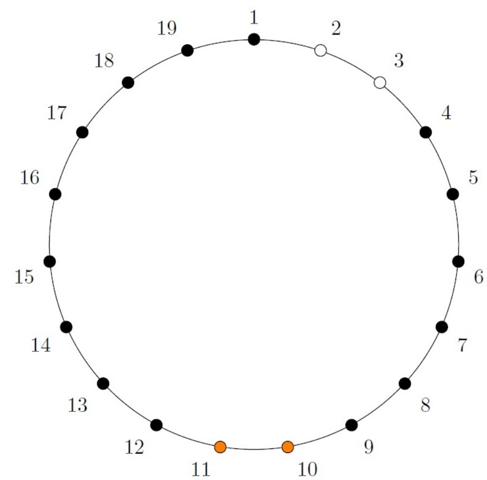

Figure 1.2: Scheme of the distribution of agents in the game (case N = 19).

1.2. Strategies

Next, we introduce a description of the social norm the community is trying to implement using group leaders. This is a symmetric strategy consisting of two phases. Furthermore, the social norm is defined so that all the players of a group will simultaneously be playing in any stage t according to the same phase. Both phases of play are described in detail below. Remember that D_t^k was the subset of players belonging to a set $k \in \{1, \ldots, N\}$ that have deviated from the social norm in any stage until stage $(t-1)$.

Phase I (also called "Cooperation Phase"):

Rule of Play:

- In Phase I, every player i $\in k$ such that $m_{\mathrm{i}}(t) = j \notin k$ (he is matched with an agent belonging to a group different from his own) plays C.
- In Phase I, every player i $\in k$ such that $m_{\mathrm{i}}(t) = j \in k$ (he is matched with an agent belonging to his same group) plays C only if the player $j \in k \wedge j \notin D_t^k$, that is, player j has never deviated before from playing the action prescribed by the social norm. Otherwise, agent i plays D.

Rule of Transition:

If a deviation from the action profile (C, C) is observed in the stage t for any of the interactions in which participated a player $\mathrm{i} \in k$ such that $m_{\mathrm{i}}(t) = j \notin k$, the entire set of players will start to play according to Phase II (or "Punishment Phase") with a probability p^k, ex-ante known for all the players in the game. This probability will be implemented through the use of a random device independent for every set of players whose realization, as already mentioned, is only watched by the group leader. It is assumed that the realizations of the random devices used for different groups have no correlation.

Phase II is defined by the following rules of play and transition:

Phase II (also called "Punishment Phase"):

Rule of Play:

- In Phase II, every player $\mathrm{i} \in k$ such that $m_{\mathrm{i}}(t) = j \notin k$ plays D.
- In Phase II, every player $\mathrm{i} \in k$ that is matched with another player (denoted by j) from the same set of agents plays C only if $j \in k \wedge j \notin D_t^k$, that is, the player j has never deviated before from playing the action prescribed by the social norm. Otherwise, agent i plays D.

Rule of Transition:

A group k playing according to Phase II will return to Phase I with probability $(1 - p^k)$, in accordance to the realizations of the random device associated to the group.

Connecting the social norm with the communication actions defined above, for group k to play according to Phase I or Cooperation Phase in stage $(t+1)$ would correspond to $A_t^{k,l} = C$, $\forall l \in \{1, 2, \ldots, k-1, k+1, \ldots, N\}$. Meanwhile, to play according to Phase II or Punishment Phase would be determined by: $A_t^{k,l} = D$, $\forall l \in \{1, \ldots, k-1, k+1, \ldots, N\}$.

1.3. Information

We now formally define the information each group leader has. We start by noticing that as a group leader sends the same signal S_t^k to all the members of his group, the group leaders of other groups can always tell the difference between a mistake made by one agent and an entire group of players following Punishment Phase or Cooperation Phase. Since, as each group is a continuum, there will always be an arbitrarily large number of members of a group k facing members of any other group l of players.

For instance, suppose that in the case $N = 19$ (shown in Figure 1.2) the set of players denoted by the number 10 (located between the orange dots) starts playing according to

Phase II in the stage $t^\star$. Since a finite number of agents belonging to set number 2 (located between the white dots in the figure) will be matched with a player from set number 10, then the group leader of 2 will know with certainty that the set 2 was playing according to Phase II in $t^\star$. We will extensively use this property in the model below.

Next, let the following vector be defined as $\forall t \geq 1$, $\forall k \in \{1, \ldots, N\}$.

$$A_t^{kk} = \left\{A_t^{1,k}, A_t^{2,k}, \ldots, A_t^{k-1,k}, A_t^{k+1,k}, \ldots, A_t^{N,k}\right\}$$

This vector includes $(N-1)$ communication actions, or equivalently, all the instructions the members of the other groups receive regarding group k in stage $(t-1)$. For reasons discussed above, this vector will be known to each group leader k.

Using this, we can define the following vectors with the history of instructions about group k until stage t.

$$H_{A,t}^{l,k} = \left\{A_1^{l,k}, A_2^{l,k}, \ldots, A_t^{l,k}\right\}$$
$$H_{A,t}^{k} = \left\{A_1^{kk}, A_2^{kk}, \ldots, A_t^{kk}\right\}$$

The first vector $(H_{A,t}^{l,k})$ are the instructions a given group leader l has sent regarding group k until stage t; while the second vector, $H_{A,t}^{k}$, corresponds to the entire history of instructions all of the other groups' players have received about group k until stage t.

In addition, we define the set of the history of the realizations of the random device of group k in every stage until stage t.

$$H_{RD,t}^{k} = \left\{RD_1^k, RD_2^k, \ldots, RD_t^k\right\}$$

Where RD_t^k is the realization of the random device of group k in stage t. The information the group leader k has at the end of stage t is, then, defined as follows:

$$I_t^k = \left\{H_{A,t-1}^{k}, H_{RD,t}^{k}\right\}$$

Meaning that each group leader knows what the realizations of his own random device have been, and the communication actions of the rest of the leaders in the past about his own group.

Chapter 2

Model without leaders

2.1. Simplified version of the model

In this section, we will study a simplified version of the model in which it will be assumed that instead of having a leadership structure, there is a perfect transmission of information within each group, so agents have perfect information about the histories of all the other agents that belong to their same group. This information includes the identity of all the players each of the agents of the group have faced in the past and in the current stage, and the profiles of actions resulting from each one of those interactions. Additionally, every player i that belong to group k knows the realization of the random device of his group.

Formally, using the sets defined in the previous section, now every player of group k knows the information that previously had to be processed by the group leaders and informed through the signal, given by:

$$I_t^k = \{H_{A,t}^k, H_{RD,t}^k, D_t^k\}$$

Players, however, still receive no direct information regarding the histories of the players belonging to groups different from their own (excluding, of course, the information they receive indirectly through the players in their own set).

The social norm to be implemented will be exactly the same as the one described above. The difference between this simplified model and the model of the previous section is just that no role is fulfil by the group leaders, and this agents don't exist in the simplified version. Therefore, the incentive problem of the group leaders, that will be discussed in the next section, is not present.

We will use this simplified version of the model both as a benchmark and to describe the incentive compatibility conditions that have to be met so that the non-leader agents have incentive to participate in the game described in the previous section.

2.2. Incentive Compatibility Conditions

The objective is to characterize restrictions in the parameters of the model (in particular, in the value of the discount factor δ) such that it is incentive compatible for the agents to play the above described symmetric strategy and cooperation (determined by stage action profile (C, C)) is sustained on the path of the equilibrium of the simplified model.

In order to define the incentive compatibility conditions, first it is necessary to define the average continuation value of the game per stage, which is defined considering for a player i $\in k$ the expected value he will obtain in interactions with players belonging to sets of players different from his own ($l \in \{1, \ldots, N\}$, such that $l \neq k$).

The above mentioned continuation value is denoted by $v\left(\frac{\mathrm{d}}{N}, \mathbf{p}, \delta\right)$, and depends on the discount factor δ (which will be assumed identical for all the agents from now on), the vector of probabilities $\mathbf{p} = \left\{p^1, \cdots p^k, p^N\right\}$ with which the different groups transition between phases, and the proportion of sets of players that have transitioned to play in Phase II to the total number of groups of players: $\left(\frac{\mathrm{d}}{N}\right)$.

The model will be solved for simplicity for the case in which the transition probabilities are the same for every group of players: $p^1 = p^2 = \ldots = p^k = \ldots = p^N = p$. The incentive compatibility condition that has to be fulfilled so that the social norm is incentive compatible for an agent that is in Phase I and is matched with a player who is playing according to Phase I as well, is the following:

$$\delta(1+g) - g \geq \left(\frac{N-1}{N}\right)\delta\left[p^2 v\left(\frac{2}{N}, p, \delta\right) + 2p(1-p)v\left(\frac{1}{N}, p, \delta\right) + (1-p)^2 v\left(\frac{0}{N}, p, \delta\right)\right] \quad (2.1)$$

It is necessary to define as well the incentive compatibility conditions that have to be fulfilled so that adhering to the social norm is also incentive compatible for a player that is in Phase II.

In this game as every group of players is a continuum, as it was mentioned, there will be an arbitrarily large number of players from a group k facing players from any of the other groups of players. This means that from the point of view of a single player, if he decides to play (C) instead of the action (D) that he is supposed to, the effect this deviation can have in stopping the contagion to the group he is facing is null. Therefore the incentives compatibility conditions below will always be met, for any value of δ.

In the case when a player Phase II is matched with a player who is playing according to Phase I, there is the following incentive compatibility condition.

$$(1+g)+\left(\frac{1}{N}\right)\frac{\delta}{(1-\delta)}\cdot 1+\frac{\delta}{(1-\delta)}\left(\frac{N-1}{N}\right)\mathbb{E}_{\mathrm{d}}\left[pv\left(\frac{\mathrm{d}+1}{N},p,\delta\right)+(1-p)\,v\left(\frac{\mathrm{d}}{N},p,\delta\right)\right]\geq 1+\frac{\delta}{(1-\delta)}\cdot 0$$
$$+\frac{\delta}{(1-\delta)}\left(\frac{N-1}{N}\right)\mathbb{E}_{\mathrm{d}}\left[pv\left(\frac{\mathrm{d}+1}{N},p,\delta\right)+(1-p)\,v\left(\frac{\mathrm{d}}{N},p,\delta\right)\right] \quad (2.2)$$

In the last expression, the expectation is defined over the variable d, that refers to the number of set of players that are in Phase II in stage t.

Lastly, the next is the incentive compatibility condition when a player in Phase II is matched with a player also in Phase II.

$$-l+\frac{\delta}{(1-\delta)}\left(\frac{N-1}{N}\right)\mathbb{E}_{\mathrm{d}}\left[v\left(\frac{\mathrm{d}}{N},p,\delta\right)\right]\geq 0+\frac{\delta}{(1-\delta)}\left(\frac{N-1}{N}\right)\mathbb{E}_{\mathrm{d}}\left[v\left(\frac{\mathrm{d}}{N},p,\delta\right)\right]$$
$$+\frac{\delta}{(1-\delta)}\left(\frac{1}{N}\right) \quad (2.3)$$

We need to prove that there are values of $\delta \in [0,1]$ such that the incentive compatibility condition for a player in Phase I is met. In order to prove that, the following *lemma* will be useful.

Lemma 1 *The average per stage continuation value: $v(\frac{\mathrm{d}}{N},\delta,p)$ is convex in $\left(\frac{\mathrm{d}}{N}\right)$, the proportion of sets of players that have transitioned to play in Phase II to the total number of sets of players, when a simple condition over the pairings along the game is met.*

Demostración. See *Appendix*. □

Now, we can prove the following *proposition*.

Proposition 2 *In the simplified game with random matching previously described there is a critical value of the discount factor $\bar{\delta}$ such that $\forall\delta\in\left[\tilde{\delta},1\right)$, to follow the social norm defined by the symmetric strategy is a subgame perfect equilibrium and every player plays (C) in every stage of the game.*

Demostración. The proof of this proposition draws heavily on *Ellison* (1994). It is necessary to prove the existence of a value of the parameter δ such that for an individual agent the incentive compatibility condition in Phase I is fulfilled (since Phase II is always met). In particular, it will be proven that a value $\tilde{\delta}$ exists such that $\forall \delta \in \left[\tilde{\delta}, 1\right)$ the incentive compatibility condition is fulfilled.

We will start by fixing the value of δ in $\bar{\delta}$, which we define as the value of the discount factor that makes the incentive compatibility condition in Phase I to be met with equality.

$$\bar{\delta} - g(1-\bar{\delta}) = \left(\frac{N-1}{N}\right)\bar{\delta}\left[p^2 v\left(\frac{2}{N}, p, \bar{\delta}\right) + 2p(1-p)v\left(\frac{1}{N}, p, \bar{\delta}\right) + (1-p)^2 v\left(\frac{0}{N}, p, \bar{\delta}\right)\right]$$

Dividing by: $(1-\bar{\delta})$

$$\frac{\bar{\delta}}{(1-\bar{\delta})} - g = \left(\frac{\bar{\delta}}{1-\bar{\delta}}\right)\left(\frac{N-1}{N}\right)\left[p^2 v\left(\frac{2}{N}, p, \bar{\delta}\right) + 2p(1-p)v\left(\frac{1}{N}, p, \bar{\delta}\right) + (1-p)^2 v\left(\frac{0}{N}, p, \bar{\delta}\right)\right]$$
$$\Longleftrightarrow g = \left(\frac{\bar{\delta}}{1-\bar{\delta}}\right)\left\{1 - \left(\frac{N-1}{N}\right)\left[p^2 v\left(\frac{2}{N}, p, \bar{\delta}\right) + 2p(1-p)v\left(\frac{1}{N}, p, \bar{\delta}\right) + (1-p)^2 v\left(\frac{0}{N}, p, \bar{\delta}\right)\right]\right\}$$

To simplify the notation from now on, instead of $v\left(\frac{\mathrm{d}}{N}, p, \bar{\delta}\right)$, it is written $v(\mathrm{d})$ and it is defined the variable α:

$$\begin{aligned}\alpha &= \left\{1 - \left(\frac{N-1}{N}\right)\left[p^2 v\left(\frac{2}{N}, p, \bar{\delta}\right) + 2p(1-p)v\left(\frac{1}{N}, p, \bar{\delta}\right) + (1-p)^2 v\left(\frac{0}{N}, p, \bar{\delta}\right)\right]\right\} \\ &= \left\{1 - \left(\frac{N-1}{N}\right)\left[p^2 v(2) + 2p(1-p)v(1) + (1-p)^2 v(0)\right]\right\}\end{aligned}$$

It is necessary to check if α is strictly positive. Lets assume this is true to see what condition over the parameters of α this would imply.

$$\begin{aligned}\alpha = 1 - \left(\frac{N-1}{N}\right)&\left[p^2 v(2) + 2p(1-p)v(1) + (1-p)^2 v(0)\right] > 0 \\ \Longrightarrow 1 > \left(\frac{N-1}{N}\right)&\left[p^2 v(2) + 2p(1-p)v(1) + (1-p)^2 v(0)\right]\end{aligned}$$

As $\left(\frac{N-1}{N}\right) < 1$, it suffices to show that $1 > \left[p^2 v(2) + 2p(1-p)v(1) + (1-p)^2 v(0)\right]$ in order to demonstrate that α is bigger than zero.

$$p^2 v\,(2) + 2p(1-p)v\,(1) + (1-p)^2 v\,(0) < 1$$
$$p^2 v(2) + 2p(1-p)v(1) + v(0) - 2pv(0) + p^2 v(0) < 1$$

Using now that, by definition, $v(0) = 1$.

$$p^2 v(2) + 2p(1-p)v(1) + \cancel{1} - 2p + p^2 < \cancel{1}$$
$$p^2 v(2) + 2pv(1) - 2p^2 v(1) + p^2 < 2p \quad / \cdot \frac{1}{p}$$
$$pv(2) + 2v(1) - 2pv(1) + p < 2$$

This last equation can be re-ordered as:

$$p\,(1 + v(2)) + (1-p)\,(2v(1)) < 2$$

As each one of the terms in parentheses is smaller than 2 (since $v(1), v(2) < 1$) the whole expression, which is the convex combination of both terms, has to be less than 2 as well. With this, is has been proven that always $\alpha > 0$.

Now using that $\alpha > 0$, it is not hard to see that:

$$\lim_{\delta \to \infty} \frac{\delta}{1-\delta} \cdot \alpha = \infty$$
$$\lim_{\delta \to 0} \frac{\delta}{1-\delta} \cdot \alpha = 0$$

So, from the relationships above and using continuity in δ, it is clear that there has to exist a $\bar{\delta} \in (0,1)$ such that:

$$\frac{\bar{\delta}}{1-\bar{\delta}} \cdot \alpha = g$$

Or equivalently,

$$\left(\frac{\bar{\delta}}{1-\bar{\delta}}\right) \left\{1 - \left(\frac{N-1}{N}\right) \left[p^2 v\left(\frac{2}{N}, p, \bar{\delta}\right) + 2p(1-p)v\left(\frac{1}{N}, p, \bar{\delta}\right) + (1-p)^2 v\left(\frac{0}{N}, p, \bar{\delta}\right)\right]\right\} = g$$

Now, we turn back our attention to the incentive compatibility condition in Phase I.

$$\delta(1+g)-g \geq \left(\frac{N-1}{N}\right)\delta\left[p^2 v\left(\frac{2}{N},p,\delta\right)+2p(1-p)v\left(\frac{1}{N},p,\delta\right)+(1-p)^2 v\left(\frac{0}{N},p,\delta\right)\right]$$

We focus in the case when this equation is satisfied with equality, which we already know exists for what was proven above.

$$\delta(1+g)-g = \left(\frac{N-1}{N}\right)\delta\left[p^2 v\left(\frac{2}{N},p,\delta\right)+2p(1-p)v\left(\frac{1}{N},p,\delta\right)+(1-p)^2 v\left(\frac{0}{N},p,\delta\right)\right]$$
$$\iff \delta(1+g)-g = \left(\frac{N-1}{N}\right)\delta\beta$$

In the last equation, it was used the variable β. Defined as:

$$\beta = \left[p^2 v\left(\frac{2}{N},p,\delta\right)+2p(1-p)v\left(\frac{1}{N},p,\delta\right)+(1-p)^2 v\left(\frac{0}{N},p,\delta\right)\right]$$

.

We notice that when the incentive compatibility condition is met with equality in Phase I, a player is indifferent in any stage between playing C or D. This means that it has to be equivalent for him to play C and receive a continuation value of $v(0)$, and play D and receive a continuation value $v(1)$. Taking a closer look at the variable β, this implies that the following has to be true:

$$\begin{aligned}\beta = \left[p^2v(2)+2p(1-p)v(1)+(1-p)^2v(0)\right] &= \left[p^2v(1)+p(1-p)v(1)+p(1-p)v(0)+(1-p)^2v(0)\right]\\ &= \left[v(1)\left(\cancel{p^2}+p-\cancel{p^2}\right)+v(0)\left(p-\cancel{p^2}+1+\cancel{p^2}-2p\right)\right]\\ &= \left[v(1)p+v(0)\left(1-p\right)\right] = \left[-p\left(v(0)-v(1)\right)+1\right]\end{aligned}$$

Replacing in the incentive compatibility condition in Phase I.

$$\delta(1+g)-g = \left(\frac{N-1}{N}\right)\delta\left[-p\left(v(0)-v(1)\right)+1\right]$$

Re-arranging the last equation:

$$
\begin{aligned}
\left(\frac{N-1}{N}\right)\delta p\left(v(0)-v(1)\right) &= \left(\frac{N-1}{N}\right)\delta-\delta(1+g)+g \\
\left(\frac{N-1}{N}\right)\delta p\left(v(0)-v(1)\right) &= \left(\frac{(N-1)\delta-N\delta}{N}\right)-\delta g+g \\
\left(\frac{N-1}{N}\right)\delta p\left(v(0)-v(1)\right) &= -\delta\left(\frac{1}{N}\right)+g(1-\delta) \\
g-\frac{\delta}{(1-\delta)}\left(\frac{1}{N}\right) &= \frac{1}{(1-\delta)}\left(\frac{N-1}{N}\right)\delta p\left(v(0)-v(1)\right)
\end{aligned}
$$

We can now replace $(v(0)-v(1))$ using the equation found in the proof of the convexity of the continuation value of the game.

$$
\begin{aligned}
g-\frac{\delta}{(1-\delta)}\left(\frac{1}{N}\right) &= \frac{1}{\cancel{(1-\delta)}}\left(\frac{N-1}{N}\right)\delta p\sum_{t=0}^{\infty}\cancel{(1-\delta)}\delta^t p^t R^t(p,N)\left[(g)\text{Prob}\left(C_{11}\right)-(l+1)\text{Prob}\left(C_{12}\right)-(1)\text{Prob}\left(C_{22}\right)\right] \\
g-\frac{\delta}{(1-\delta)}\left(\frac{1}{N}\right) &= \left(\frac{N-1}{N}\right)\sum_{t=0}^{\infty}(\delta p)^{t+1}R^t(p,N)\left[(g)\text{Prob}\left(C_{11}\right)-(l+1)\text{Prob}\left(C_{12}\right)-(1)\text{Prob}\left(C_{22}\right)\right]
\end{aligned}
$$

Where C_{nm} makes reference to the probability player n is playing according to Phase I and the player m he meets is playing according to Phase II.

The right hand side of the expression depends only on δ and p. When $p=1$, then $R^t(1,N)=1$, $\forall N\in\mathbb{N}$. Fixing $\delta=\bar{\delta}$, then we know the Phase I condition is met with equality.

By choosing $p(\delta)$ such that $t\left[R(p(\delta),N)\delta p(\delta)\right]'+R(p(\delta),N)(\delta p(\delta))'=0$, then the right hand side is constant $\forall\delta\in\left[\bar{\delta},1\right]$. This implies that the incentive compatibility condition is satisfied for every δ within this interval, and this ends the proof of existence.

$\square$

Chapter 3

Model with Group Leaders

3.1. The incentive problem of the group leaders

We return now to the model, presented in Chapter II, with the presence of group leaders. We will start by briefly describing the incentive problems group leaders face.

The incentive problems of the leaders in this case is generated by private information. Only the group leader of group k knows what the realizations of the random device of his group have been. Therefore, he can be naturally tempted to lie about the result of the random device and give dishonest instructions to the members of his group. Remember that we have assumed that group leaders payments are correlated to the payments of the members of his group so, in this particular model, the leader of group k when lying is cheating the members of the other groups of players in general, and in particular, the other group leaders.

For instance, lets suppose the leader of group k observes that group l goes into Phase II and that the realization of the random device tells him to signal the members of his group to transition to Phase II as well. In this case, to transition to Phase II might be a very costly action, so the temptation to lie and remain in Phase I could be strong. However, this could make the existence of Phase II an empty threat in practice, causing the social norm that sustains cooperation in equilibrium to collapse.

We can describe the problem group leaders face, then, as a symmetric principal-agent problem. Since every group leader can be considered an agent of the rest of the leaders (the principal) with the task of honestly applying the social norm rules. Although every leader is an agent of the whole set of the rest of the group leaders, we will be interested only in a solution of this principal-agent problem that is bilateral (i.e., in which every pair of leaders supervise each other's behaviour independently).

In order for incentives to be aligned, the leaders must be able to detect and punish a leader of a different group when he cheats. In this case, since defections from the social norm will only be noticeable after a sufficiently large number of stages have elapsed, the supervision of the behaviour of the other leaders will have to be done using statistics of the past behaviour of each of the other leaders.

Notice also that we must only worry about the stages of the game after a deviation occurs. In what comes next, the discussion will always be making reference only to these stages of the game.

In this symmetric game, we can conceptually separate the strategy the group leader k follows in two parts: a review strategy used to supervise $\sigma^k = \left\{\sigma^k_{t+1}(I^k_t)\right\}_{t=0}^{\infty}$; and a strategy leader k uses an *agent* of the rest of the leaders in the game, $\tau^k = \left\{\tau^k_{t+1}(I^k_t)\right\}_{t=0}^{\infty}$. For exposition purposes, we will refer to the latter as *communication strategy* since it consists of the orders, conditionals in the history of the game, that the group leader communicates to his group of players. Lets notice that, although defined here separately, the review and communications strategies a group leader follows are interdependent. In particular, as it will be shown below, a review strategy will define the communication strategy used by the leader doing the supervision (acting as principal).

Since we are looking to characterize a bilateral review strategy, the problem of each group leader regarding the supervision of the rest of the leaders can be decomposed in $(N-1)$ review problems. As the behaviour of a third leader is irrelevant in a bilateral review problem, then the review strategy used by group leader k against group leader l should only depend on the past actions of group l against group k: $H^{l,k}_{A,t}$. In what comes next, we will refer to the review strategy of the group leader k about the group l as $\sigma^{k,l} = \left\{\sigma^{k,l}_{t+1}(I^k_t)\right\}_{t=0}^{\infty}$, and analogously to the communication strategy of the group leader k about the group l as $\tau^{k,l} = \left\{\tau^{k,l}_{t+1}(I^k_t)\right\}_{t=0}^{\infty}$

In order to explain what comes next, it is useful to enunciate the following *lemma*.

Lemma 3 *The incentives group leaders have for deviating from the social norm in stage t depends on the parameters of the model ($\boldsymbol{p}$, l, g) and are unaffected by the number of group of players in Punishment Phase in the previous stage of the game (denoted by d_{t-1}), if $\mathrm{d}_{t-1} \geq 1$.*

Demostración. See *Appendix.* □

The meaning of the last *lemma* is that no matter the history of the game after the first deviation, if there was at least one group of agents playing according to Punishment Phase in the previous stage $(t-1)$, then:

- The number of group of players in Phase II in stage t is an independent event respect to the number of group of players in Phase II in the stage $(t-1)$ of the game.
- Because of the latter, the incentives the group leaders have to cheat will always be in one direction. That is, group leaders will want to lie about the realization of their random device either always saying $A_t^{k,l} = D$, $\forall l, t$; or $A_t^{k,l} = C$, $\forall l, t$.

3.2. Symmetric equilibrium solution for the model with group leaders

We will characterize a symmetric equilibrium for the group leaders, meaning that all group leaders will be using the same bilateral review strategy $\sigma^{k,l}$. Where group leader k is doing the supervision and group leader l is being supervised. This review strategy will consist of a trial period lasting R stages of the game, during which, the leader of the group k plays following his best response as an agent, conditional in the history of the game, which will be denoted as $\tilde{\tau}_{t+1}^{k,l}(I_t^k)$.

At the end of the trial period, the leader of group k will compare the behaviour of group l with the behaviour a group playing exactly according to the social norm would have had during the trial period. For that, we will define for group l, its state in the stage t as $C_t^l \in \{0,1\}$. If the communication action of group l towards group k in stage t is (D), then its state will take a value of 0; otherwise taking a value of 1.

In the last paragraph we are using *Lemma* 3, since as the deviations are always in one direction independently of what happened in the previous stage, it is appropriate to build an statistic that just counts the number of times group leaders have said either (D) or (C). It is important to state that the actual direction in which the incentives to cheat are is not important for the demonstration below, although the state of the group was defined in this case assuming a configuration of the game in which the incentives for all the leaders are to cheat by falsely declaring $A_t^{k,l} = D$.

We will define a particular communication strategy $\hat{\tau}$, defined such that the leader using this strategy never lies about the realization of his random device.

$$\hat{\tau}_{t+1}^{l,k}(I_t^l) = \begin{cases} C, & \text{if } RD_t^l = C \\ D, & \text{if } RD_t^l = D \end{cases}$$

For this particular communication strategy then C_t^l would be a random variable with independent realizations and, moreover, after R stages had elapsed after the first deviation, we would have that:

$$\mathbb{E}\left[\sum_{t=1}^{R} C_t^l\right] = \mathbb{E}\left[S_R^l\right] = (1-p)R = qR$$

where S_t^l will be referred to as the accumulated state of the group l after R stages, and q is the complement of the probability p of being in Punishment Phase in the social norm if at least one group was in Punishment Phase in the previous stage.

At the end of the trial period, the accumulated state of the group being controlled (S_R^l) is compared against a threshold:

$$S = qR - B$$

Where qR is the expected value of the accumulated state for every group after R stages of the game and B is a variable that allows to adjust the *severity* of the threshold. If the group l passes the trial period then a new trial period with the same characteristics described above will start. On the contrary, if the threshold is not met, then a penalty period lasting M stages will start, during which the leader of the group k will tell the agents of his group to play the Nash Equilibrium of the stage game (in this case, to play D) when facing members of group l.

In order to define $\sigma^{k,l}$ we need to establish $TI_{k,l}$ and $TII_{k,l}$, disjoint subsets to which are assigned the stages of the game (from $t = 1$ on), according to the history of play between k and l. We will start by naming a cycle of the game as the sum of a trial period and its potential following penalty period. Therefore, the cycles of the game (including its trial period and possible penalty period) can be numbered by a variable $n \in \mathbb{N}$.

Lets denote the accumulated state for group k in cycle n as $S_R^k(n)$ and define a binary variable regarding the n-th cycle penalty period as $P_{k,l}(n)$. Using this:

$$P_{k,l}(n) = \begin{cases} 0 & \text{, if } S_R^k(n) \geq S \;\wedge\; S_R^l(n) \geq S \\ 1 & , \;\sim \end{cases}$$

Now, we will define a variable Q that will let us count how many penalty periods have been before the n-th cycle.

$$Q(k) = \sum_{n=1}^{n-1} P_{k,l}(n)$$

Lastly, lets define sets that will gathered the stages groups k and l have been punishing each other (in penalty period).

$\forall n \in \mathbb{N}$,

$$R_{k,l}(n) = \begin{cases} \{nR + QM + 1, nR + QM + 2, \ldots, nR + (Q+1)M\} & \text{, if } P_{k,l}(n) = 1 \\ \emptyset & \text{, if } P_{k,l}(n) = 0 \end{cases}$$

Using these, we can formally define the sets $TI_{k,l}$ and $TII_{k,l}$.

$$\begin{aligned} T &= \{1, 2, \ldots, \infty\} \\ TII_{k,l} &= \{t : \exists n \quad \text{s.t. } t \in R_{k,l}(n)\} \\ TI_{k,l} &= T \setminus TII_{k,l} \end{aligned}$$

Using the latter, we define $\sigma^{k,l}$.

$$\sigma^{k,l}_{t+1}(I^k_t) = \begin{cases} \tilde{\tau}^{k,l}_t(I^k_t), & \text{if } t \in TI \\ D, & \text{if } t \in TII \end{cases}$$

We can consider that in the case the leader of k deviates from this strategy in a way not included in the mirror review by group l to group k (for instance, that the leader of k starts a penalty period after the threshold is met), then it would be triggered a penalty period from l to k of infinite duration.

We will prove that under an aptly designed review strategy $\sigma^{k,l}$, when all the leaders are using $\sigma^{k,l}$ to bilaterally and simultaneously control the other leaders, under any optimum symmetric strategy $\hat{\tau}^{k,l}$ that group leaders have as an agent, the expected discounted equilibrium utility of every leader can be made arbitrarily close to the Pareto efficient stage outcome for the group leaders ($\hat{v}$), defined as the expected stage payment a group leader would obtain if every leaders in the game was using $\hat{\tau}$ as an agent of the rest. We will use in what comes next the following definitions.

Definition 1*: a symmetric equilibrium of the game is characterized by a symmetric equilibrium strategy profile $S^* = \{s^*_1, \ldots, s^*_N\}$, where $\{s^*_1 = s^*_2 \ldots = s^*_N = s^*\}$, with s^* consisting of a pair $\{\sigma^*, \tau^*\}$ of a review strategy σ^* and a communication strategy τ^* such that, when every group leader is using s^**:

- τ^* *is an optimal communication strategy against review strategy σ^** being used by the rest of the leaders .
- σ^* *specifies the use of communication strategy τ^* during the trial period.*

Definition 2*: a symmetric equilibrium is said to be "robust", if after an unilateral deviation from the on-the-path action profile in any stage of the game, the on-the-path stage action* *profile is expected to be restored after a finite number of stages of the games have elapsed.*

The latter definition, says that in this case a "robust" equilibrium strategy profile will have the property that after a first deviation from the on-the path stage profile (C, C) happened, possibly causing some groups of players to transition to Phase II, after a finite number of stages of the game have elapsed, we will see again that in every interaction is being played the on-the-path stage profile (C, C). In other words, cooperation will be eventually restored.

Theorem 4 *In the model with group leaders it is possible to define a bilateral review strategy σ, defined by the pair (R, M), and by the use of communication strategy $\tilde{\tau}$ during the trial period; with R being is the finite number of stages the trial period lasts, M being the finite number of stages the penalty period lasts, and $\tilde{\tau}$ being the optimal communication strategy against σ, such that:*

- *$\{\sigma, \tilde{\tau}\}$ constitute a symmetric equilibrium strategy profile.*
- *The symmetric equilibrium characterized by $\{\sigma, \tilde{\tau}\}$ is robust.*
- *For both group leaders their equilibrium discounted payoffs are arbitrarily close to the Pareto efficient stage payoff.*

Sketch of the Proof

The idea of the proof is that by smartly picking the parameters R, M and B of the bilateral review strategy, the ex-ante expected utility of a group leader being supervised by the review strategy can be made arbitrarily close to the Pareto efficient stage utility when he is using a communication strategy that is optimal against the review strategy.

Indeed, for a group leader to play honestly following to the letter the social norm is always an available option (that we will denote as "honest" strategy below: $\hat{\tau}$), and the discounted expected utility that this option brings will be a lower bound in the possible utility of the group leader. We will also show that when the discount factor is close to unity, if the probability of failing the trial period when using $\hat{\tau}$ (that is denoted by $\hat{\phi}$) can be made close to zero, then the expected discounted utility would be approximately the Pareto efficient stage payoff, $\hat{v}$.

Using the latter, and making also use of Chevyshev's inequality we can show that if the trial period lasts for long enough, then the probability of failing the trial period $\hat{\phi}$ when using the honest strategy can me made arbitrarily small, so that the lower bound in the utility of the group leader can be made arbitrarily close to $\hat{v}$ for a value of the discount factor that is high enough.

Later on, we will obtain an upper bound for the expected discounted utility when the group leader is using any possible communication strategy τ. Using this value, when $\delta = 1$

we can establish a value of ϕ that assures that the upper bound of the utility is at least at a distance ε of $\hat{v}$.

Using these expression relating to the upper and lower bounds in the utilities, it can finally be proved that when the group leader is using his best communication strategy $\tilde{\tau}$ against a review strategy with parameters R, $M(R)$, and $B(R)$ (defined as in *Radner,1985*), for values of R and δ of at least R_ε and δ_ε, respectively, and for any $\varepsilon > 0$ that is sufficiently small, the corresponding probability of failing the trial period $\tilde{\phi}$ for the group leader will be smaller than ε and the expected discounted utility will be at a distance of at most $\varepsilon/4$ of $\hat{v}$, the Pareto efficient stage payoff.

DEMOSTRACIÓN. It will be proved for the case in which $p_1 = p_2 = \ldots = p_N = p$, $\forall k$. Without loosing generality, and using the *Lemma* 3, we will assume that according to the parameters of the game the group leaders have incentives to cheat by always declaring to play Punishment Phase (there would not be any changes in the essential parts of the argument, had the incentives to cheat be in the opposite direction).

To increase the clarity of the proof, we will prove the *Theorem* considering the interactions between the leaders of only two groups (from groups 1 and 2, respectively). As the game is completely symmetric regarding the principal-agent relations between any pair of leaders in the game, this implies that the *Theorem* 4 holds for any pair of group leaders in the game.

To make matters simpler to understand, we will consider the leader of group one to be the acting as a principal (following the review strategy $\sigma_{1,2}$), and the leader of the group number two to be acting as an agent that follows communication strategy $\tau_{2,1}$. Although, according to the model, simultaneously there is an agent-principal relationship between these same leaders in which leader 1 acts as the agent and leader 2 acts as the principal. As it was already stated, the review strategy $\sigma_{1,2}$ consists of a trial period lasting R stages of the game, during which, the leader that is acting as principal (leader of group 1) plays following a strategy $\tau_{1,2}$. At the end of the trial period, the accumulated state of the group being controlled (S_R^2) will be compared against a threshold:

$$S = qR - B$$

We want to proof the existence of a symmetric equilibrium, for that reason in what follows when talking of $\sigma_{1,2}$ it will be understood that the communication strategy being played by group leader 1 during the trial period is $\tau_{1,2} = \tau_{2,1} = \tilde{\tau}$. Furthermore, if we were to take a look over the supervision group leader 2 does over group leader 1, he would be also using $\sigma_{2,1} = \sigma_{1,2} = \sigma$. We want to characterize a find a pair (R, M) for σ such that when $\tau_{2,1} = \tilde{\tau}$ the *Theorem* 4 is true.

We will denote in what follows $V_t^2(\sigma_{1,2}, \tau_{2,1})$ as the expected stage payoff for leader 2 in the trial phase. This value depends both on the review strategy $\sigma_{1,2}$ and the strategy followed by the leader acting as agent: $\tau_{2,1}$.

We will latter use $\hat{v}$, defined as the expected stage payoff for the leader acting as agent when he is is playing using communication strategy $\hat{\tau}$ (defined above) in which he plays strictly following the social norm, and the other leader is using the review strategy $\sigma_{1,2} = \hat{\sigma}$, which means that leader of group 1 is also employing $\hat{\tau}$ during the trial period.

Moreover, when a leader has failed the trial phase, as the leader acting as principal will punish his group by playing the Nash Equilibrium action of the stage game the best thing he can do is to tell the players of his group to also play the NE of the stage game (in this case, to play D). We will call the expected stage payment of the NE for the leader acting as agent as v^*.

Using these definitions and dropping for simplicity in the exposition the group indexes, the discounted expected payment of group 2 leader is defined by the following expression.

$$\begin{aligned} v(\delta) = (1-\delta)\sum_{t=1}^{R}\delta^{t-1}V_t(\sigma,\tau) + \phi\left[\delta^R(1-\delta^M)v^* + \delta^{R+M}v(\delta)\right] + \\ (1-\phi)\left[\phi\left(\delta^R(1-\delta^M)v^* + \delta^{R+M}v(\delta)\right) + (1-\phi)\delta^R v(\delta)\right] \\ \iff v(\delta)\left(1-\phi\delta^{R+M} - (1-\phi)\phi\delta^{R+M} - (1-\phi)^2\delta^R\right) \\ = (1-\delta)\sum_{t=1}^{R}\delta^{t-1}V_t(\sigma,\tau) + \phi\delta^R(1-\delta^M)v^* + (1-\phi)\phi\delta^R(1-\delta^M)v^* \end{aligned}$$

Re-arranging:

$$v(\delta) = \frac{(1-\delta)\sum_{t=1}^{R}\delta^{t-1}V_t(\sigma,\tau) + \phi\delta^R(1-\delta^M)(2-\delta^M)v^*}{1-\phi\delta^{R+M}(2-\phi) - (1-\phi)^2\delta^R}$$

Where ϕ is symmetric probability of not meeting the threshold S.

If both leaders when in trial period are using communication strategy $\hat{\tau}$, the expected discounted utility will correspond to the following value.

$$\hat{v}(\delta) = \frac{(1-\delta^R)\hat{v} + \hat{\phi}\delta^R(1-\delta^M)(2-\delta^M)v^*}{1-\hat{\phi}\delta^{R+M}(2-\hat{\phi}) - (1-\hat{\phi})^2\delta^R}$$

In the last equation, $\hat{\phi}$ is the probability of failing the trial phase, given that the strategy $\hat{\tau}$ was used, and remembering that $V_t(\hat{\sigma}, \hat{\tau}) = \hat{v}$.

Using *l'Hôpital Rule* it can be calculated the limit of the last expression when $\delta \to 1$.

$$\lim_{\delta \to 1} \hat{v}(\delta) = \frac{R\hat{v} + \hat{\phi} M v^*}{R + \hat{\phi}(2 - \hat{\phi})M}$$

This value has two important connotations. First, it gives as a lower bound on the discounted expected payment an agent can receive when using any strategy τ, as $\hat{\tau}$ is always available. Second, it can be noticed that if we could make $\hat{\phi}$ to be close to zero, then $\hat{v}(\delta)$ would tend to $\hat{v}$ for values of δ close to one.

The next step is to obtain an upper bound to the expected average payment of an agent playing his best strategy, which will be denoted as $\tilde{\tau}$.

Lets start by defining $\bar{c}(\tau) = \mathbb{E}\left[\sum_{t=1}^{R} \frac{C_t(\tau)}{R}\right]$, the expected average value over the trial period of the group state. As $C_t \in (0, 1)$, $\bar{c}$ is also bounded between zero and one. We also define $\hat{c} = \mathbb{E}\left[\sum_{t=1}^{R} \frac{C_t(\hat{\tau})}{R}\right] = (1-p) = q$. As we have assumed that the leaders have incentives to falsely declare Punishment Phase, then in any strategy $\tau \neq \hat{\tau}$, the expected value of the average group state should be necessarily lower than q. Using this, we can argue that there always exists a positive number K such that:

$$\begin{gathered} V_t(\sigma, \tau) - \hat{v} + K\left(\bar{c}(\tau) - \hat{c}\right) \leq 0 \\ \iff V_t(\sigma, \tau) \leq \hat{v} + K\left(\hat{c} - \bar{c}(\tau)\right) \end{gathered}$$

Replacing this in the expression for the discounted expected payment of group leader 2, we have that for any τ.

$$v(\delta) \leq \frac{(1-\delta^R)\hat{v} + (1-\delta)K\sum_{t=1}^{R}\delta^{t-1}(\hat{c} - \bar{c}(\tau)) + \phi\delta^R(1-\delta^M)(2-\delta^M)v^*}{1 - \phi\delta^{R+M}(2-\phi) - (1-\phi)^2\delta^R}$$

Following *Radner (1985)*, we can now define a function $f(\cdot)$:

$$f(\delta, R) = \sum_{t=1}^{R}(1 - \delta^{t-1}) = R - \frac{1-\delta^R}{1-\delta}$$

Such that the following will be always true.

$$|\sum_{t=1}^{R} \delta^{t-1}(\hat{c} - \bar{c}) - \sum_{t=1}^{R} (\hat{c} - \bar{c})| = |\sum_{t=1}^{R} (1 - \delta^{t-1})(\hat{c} - \bar{c})| \leq f(\delta, R)$$

Using this last relation, we can find an useful bound to the term $\sum_{t=1}^{R} \delta^{t-1}(\hat{c} - \bar{c})$.

$$\Rightarrow \sum_{t=1}^{R} \delta^{t-1}(\hat{c} - \bar{c}) \leq f(\delta, R) + \sum_{t=1}^{R} (\hat{c} - \bar{c}) = f(\delta, R) + R\hat{c} - \mathbb{E}\left[S_R(\tau)\right]$$

We can now use the definition of ϕ as the probability that $S_R(\tau)$ is smaller than the threshold.

$$\begin{aligned} \phi = \mathbb{P}\left(S_R(\tau) < Rq - B\right) = F(Rq - B) \\ \iff (1 - \phi) = 1 - F(Rq - B) \end{aligned}$$

On the other hand, by Markov's inequality.

$$(1 - \phi) \leq \frac{\mathbb{E}\left[S_R(\tau)\right]}{Rq - B} \iff (1 - \phi)(Rq - B) \leq \mathbb{E}\left[S_R(\tau)\right]$$

Using that $q = \hat{c}$, and going back to replace.

$$R\hat{c} - \mathbb{E}\left[S_R(\tau)\right] \leq R\hat{c} - (1 - \phi)(R\hat{c} - B) = R\hat{c} - (1 - \phi)R\hat{c} + (1 - \phi)B = \phi R\hat{c} + (1 - \phi)B$$

Now, using the latter and replacing.

$$\rightarrow R\hat{c} - \mathbb{E}\left[S_R\right] + f(\delta, R) \leq \phi R\hat{c} + (1 - \phi)B + f(\delta, R)$$

Using the last derivation, we can define an upper-bound to the value of $v(\delta)$, which we will call $v_0(\delta)$.

$$v_0(\delta) = \frac{(1 - \delta^R)\hat{v} + (1 - \delta)K\left[\phi R\hat{c} + (1 - \phi)B + f(\delta, R)\right] + \phi\delta^R(1 - \delta^M)(2 - \delta^M)v^*}{1 - \phi\delta^{R+M}(2 - \phi) - (1 - \phi)^2\delta^R}$$

This is the moment to give specific values for B and M. We will conveniently use the functional forms proposed by *Radner (1985)*.

$$B = R^{\rho}$$

$$M = \mu R$$

As C_t takes value of one with probability q and a value of zero with probability $(1-q)$, then the variance of C_t conditional in the use of $\hat{\tau}$ is:

$$\hat{\kappa} = Var(C_t \mid \hat{\tau}) = q(1-q)$$

And by Chevyshev's inequality:

$$\hat{\phi} \leq \frac{R\hat{\kappa}}{R^{2\rho}} = \frac{\hat{\kappa}}{R^{2\rho-1}}$$

This means that choosing a value of R big enough, we can get $\hat{\phi}$ arbitrarily close to zero. Remembering that $\lim\limits_{\delta\to 1} \hat{v}(\delta) = \hat{v}$ when $\hat{\phi} \to 0$, then, for any ε and values of R and δ chosen to be big enough: R_{ε} and δ_{ε}.

$$\hat{v}(\delta) > \hat{v} - \varepsilon \tag{3.1}$$

$\forall R \geq R_{\varepsilon}$ and $\delta \geq \delta_{\varepsilon}$.

Using again *l'Hôpital Rule*, we can obtain the value of $v_0(1)$

$$v_0(1) = \frac{\hat{v} + K\left[\phi\hat{c} + (1-\phi)R^{\rho-1}\right] + \phi(\mu-2)v^*}{1 + 2\phi\mu - \mu\phi^2}$$

Let define a variable η such that: $0 < \eta < \hat{v} - v^*$

Next, we want to find a value for ϕ that assures that:

$$v_0(1) \leq \hat{v} - \varepsilon' \tag{3.2}$$

For any $\varepsilon' < \eta/2$.

Simply replacing the expression of $v_0(1)$ in the condition, we can find such a value for ϕ.

$$\phi \geq \frac{KR^{\rho-1} + \varepsilon'}{KR^{\rho-1} - K\hat{c} + \mu((2-\phi)\hat{v} - v^* - (2-\phi)\varepsilon') + 2v^*}$$

Since $\phi \in (0,1)$, it suffices a ϕ define as below (where ϕ in the RHS has been replaced for the values in its domain such that they minimize the value of the denominator) so that the condition is met.

$$\phi \geq \frac{KR^{\rho-1} + \varepsilon'}{KR^{\rho-1} - K\hat{c} + \mu(\hat{v} - v^* - 2\varepsilon') + 2v^*} \tag{3.3}$$

In order that the denominator is positive, we fix μ in the following way.

$$\mu \geq \frac{K\hat{c}}{\hat{v} - v^* - \eta}$$

For what comes next, lets define $\tilde{v}(R,\delta)$ as the value of the expected maximum discounted utility for the group leader 2 when the trial period lasts R stages and the discount factor is equal to δ, and $\tilde{\phi}(R,\delta)$ as the corresponding probability of failing the trial period.

Proposition 5 *$\forall\varepsilon$ s.t. $0 < \varepsilon < 2\eta$ there exists R_ε and $\delta_\varepsilon < 1$ such that $\forall\delta \in [\delta_\varepsilon, 1)$:*

- $\tilde{v}(R_\varepsilon,\delta) \geq \hat{v} - \varepsilon/4$
- $\tilde{\phi}(R_\varepsilon,\delta) < \varepsilon$

Let R_ε a value of R that satisfies $R_\varepsilon \geq R_{\frac{\varepsilon}{4}}$ in (3,1) (so that $\hat{v}(\delta)$ is close enough to $\hat{v}$) and that the RHS of (3,3) is smaller than ε. Also lets set $\varepsilon' = \varepsilon/2$

In addition, lets suppose δ_ε be a value of δ such that $\delta_\varepsilon \geq \delta_{\frac{\varepsilon}{4}}$ in (3,1) and that for R_ε and $\delta \in [\delta_\varepsilon, 1)$ the following expression is fulfilled;

$$\mid v_0(\delta) - v_0(1) \mid \leq \frac{\varepsilon}{4} \tag{3.4}$$

Since $\tilde{v}(R_\varepsilon,\delta) \geq \hat{v}(\delta)$, $\forall\delta \in [\delta_\varepsilon, 1)$, then:

$$\tilde{v}(R_\varepsilon,\delta) \geq \hat{v}(\delta) \geq \hat{v} - \varepsilon/4$$

The last inequality implies that the first part of the *proposition* is true.

Now, lets suppose that $\delta > \delta_\varepsilon$ and $\tilde{\phi} \geq \varepsilon$. Then (3,3) and (3,2) would imply that:

$$v_0(1) \leq \hat{v} - \varepsilon/2$$

Which together with (3,4) necessarily means that:

$$v_0(\delta) \leq \hat{v} - \frac{\varepsilon}{4}$$

But $\tilde{v}(\delta) \leq v_0(\delta)$.

$$\rightarrow \tilde{v}(R_\varepsilon, \delta) \leq \hat{v} - \frac{\varepsilon}{4}$$

This is a contradiction and finishes the proof of the *proposition*.

Using this last result, using the symmetry of the game, we now can generalize and conclude that if a group leader is using a review strategy to supervise bilaterally the behaviour of all the group leaders in the game, then in all of this symmetric relationships, for a value of δ that is high enough, then the discounted utility of the group leaders will be arbitrarily close to the Pareto efficient result.

One last remaining question is whether, once there was a first deviation, the symmetric strategies being used by the group leaders would allow in expected value for a sufficiently large number of stages with at least one group playing according to Punishment Phase, such that the review strategies that can implement the Pareto efficient outcome are actually feasible to be adopted.

Lets denote the period of time in the game from the stage when the first deviation happens to the first stage when there is no longer any group of players in Punishment Phase as the group-war period.

Proposition 6 *For any duration in number of stages of the trial period of the review strategy, R, can always be found a value p of the probability for each group of staying in Punishment Phase such that the expected number of stages that the group-war period will last is at least of R periods, when the communication strategy being used in equilibrium $\tilde{\tau}$ is close to $\hat{\tau}$.*

DEMOSTRACIÓN. The expected number of stages the group-war period is going to last is equal to the expected number of stages until all the groups return to Phase I. Lets start by making the supposition that every group is following the strategy $\hat{\tau}$. Then, as the probability for every group of remaining in Phase II is exactly p, the probability that every group simultaneously goes back to Phase I is $(1-p)^N$ and the expected number of stages $T(p)$ until that event happens is:

$$T(\hat{\tau}) = \frac{1}{(1-p)^N}$$

Now, for every duration R of the trial phase, as $T(\hat{\tau})$ is increasing in p we can always find a value p such that $T(\hat{\tau}) \geq R$.

It is important to notice that in for a fix δ, according to equation (2,1), if for a value p_0 of p the incentive compatibility condition was met, then with a value $p_0 < p_1 < 1$ the condition will be equally fulfilled. As incentive compatibility condition in Phase II is always met, the "growth" in the parameter p would have no effect in the incentive compatibility condition of non-leader players.

More generally, for the optimum communication strategy $\tilde{\tau}$ we can think of two different cases. The first one, is if $\tilde{\tau}$ does not generate correlation between (d_t) and (d_{t+s}), $\forall s > 0$. In this case, the problem is stationary and we can call the probability with which any group goes back into Phase I as $\tilde{p}$. It has to be then that for values of p that are high enough, $\tilde{p}$ can be made big enough that $T(\tilde{\tau}) \geq R$.

In the case when the optimum communication strategy $\tilde{\tau}$ does generate correlation, it is (see proof of *Lemma 2*) proven that agent will always want to cheat by falsely declaring Phase II instead of Phase I. Therefore, although in this case the probability is not stationary, in any stage of the game it has to be that the probability is *at least* of value p. The latter means that in this case also $T(\tilde{\tau}) \geq T(\hat{\tau})$. $\square$

Finally, we can notice that the expected number of stages the group-war period will last $T(p, \tilde{\tau})$ will be finite. This means that after a finite number of stages, we should expect that every group of players in the game is back into Phase I. Therefore, the symmetric equilibrium defined by $\{\sigma, \tilde{\tau}\}$ is robust.

With this, the proof of the *Theorem* 4 ends.

$\square$

Conclusion

We have presented a repeated game model of inter-group interaction in which group leaders are central so that the common social norm can be implemented.

Our main contributions have been to explicitly incorporating leadership into a social norm model, and to show how a simple structure of leadership can sustain an equilibrium in which opportunistic behaviour on the part of the group leaders is controlled and inter-group relations remain cooperative on the path of the equilibrium.

There are several interesting ways in which our work could be extended. We haven't tackled in this work the internal incentive problems within each group of players: for instance, considering that a member of the group could cheat his own group by not declaring the truth about his past interactions with members from different groups; or supposing that the group leaders receive a noisy signal about their group members behaviour.

Also, we have used a very simple structure of leadership that could be, undoubtedly, perfected. We have considered an structure of leadership that is static but, in most cases in reality, the nature of leadership tends to be dynamic and there is a finite period of time over which a leader can hope to be the group's decision maker.

Finally, we have used a continuum of players in each group. However, a distinctive feature of inter-group relations is that the size different groups have, usually, influences their interactions. Therefore, a model with a finite a possibly different number of players per group would be interesting to study.

Bibliography

[1] Daron Acemoglu and Matthew O Jackson. History, expectations, and leadership in the evolution of social norms. Technical report, National Bureau of Economic Research, 2011.

[2] Avinash Dixit. Trade expansion and contract enforcement. *Journal of Political Economy*, 111(6):1293–1317, 2003.

[3] Avinash K Dixit. *Lawlessness and economics: alternative modes of governance*. Princeton University Press, 2007.

[4] Glenn Ellison. Cooperation in the prisoner's dilemma with anonymous random matching. *The Review of Economic Studies*, 61(3):567–588, 1994.

[5] Jon Elster. Social norms and economic theory. *The Journal of Economic Perspectives*, pages 99–117, 1989.

[6] Juan F Escobar. Cooperation and self-governance in heterogeneous communities. Technical report, 2008.

[7] James D Fearon and David D Laitin. Explaining interethnic cooperation. *American political science review*, pages 715–735, 1996.

[8] Clifford Geertz. *The interpretation of cultures: Selected essays*, volume 5019. Basic books, 1973.

[9] Parikshit Ghosh and Debraj Ray. Cooperation in community interaction without information flows. *The Review of Economic Studies*, 63(3):491–519, 1996.

[10] Avner Greif. Contract enforceability and economic institutions in early trade: The maghribi traders' coalition. *The American economic review*, pages 525–548, 1993.

[11] Michihiro Kandori. Social norms and community enforcement. *The Review of Economic Studies*, 59(1):63–80, 1992.

[12] Rachel E Kranton. Reciprocal exchange: a self-sustaining system. *The American Economic Review*, pages 830–851, 1996.

[13] Paul R Milgrom, Douglass C North, et al. The role of institutions in the revival of trade:

The law merchant, private judges, and the champagne fairs. *Economics & Politics*, 2(1):1–23, 1990.

[14] Masahiro Okuno-Fujiwara and Andrew Postlewaite. Social norms and random matching games. *Games and Economic behavior*, 9(1):79–109, 1995.

[15] Roy Radner. Monitoring cooperative agreements in a repeated principal-agent relationship. *Econometrica: Journal of the Econometric Society*, pages 1127–1148, 1981.

[16] Roy Radner. Repeated principal-agent games with discounting. *Econometrica: Journal of the Econometric Society*, pages 1173–1198, 1985.

[17] Roy Radner. Repeated partnership games with imperfect monitoring and no discounting. *The Review of Economic Studies*, 53(1):43–57, 1986.

[18] Satoru Takahashi. Community enforcement when players observe partners' past play. *Journal of Economic Theory*, 145(1):42–62, 2010.

Appendix A

Derivation of Incentive Compatibility Condition in Phase I

In order that a player belonging to a set currently playing according to Phase I does not want to deviate and play (D) instead of (C), it is sufficient that the following condition is fulfilled.

$$(1+g)+\left(\frac{N-1}{N}\right)\frac{\delta}{(1-\delta)}\cdot\left[p^2 v\left(\frac{2}{N},p,\delta\right)+2p\,(1-p)\,v\left(\frac{1}{N},p,\delta\right)+(1-p)^2\,v\left(\frac{0}{N},p,\delta\right)\right] + \left(\frac{1}{N}\right)\frac{\delta}{(1-\delta)}\cdot 0 \leq \frac{1}{(1-\delta)}$$

The first term on the left hand side (LHS) of this expression corresponds to the immediate gain from deviating; the second term is the continuation value with respect to the sets of players in the game different from his own, which depends on the probability $p(\delta)$ that determines how likely the sets of players that have observed the defection will now transition to Phase II. Finally, the third term on the LHS is the continuation value for the player that deviates with respect to its own set of players. As his deviation is perfectly observed by these players, he will always be punished by them with certainty, and therefore, will receive zero from these interactions.

Multiplying this inequality by $(1-\delta)$.

$$(1+g)\,(1-\delta)+\left(\frac{N-1}{N}\right)\delta\cdot\left[p^2 v\left(\frac{2}{N},p,\delta\right)+2p\,(1-p)\,v\left(\frac{1}{N},p,\delta\right)+(1-p)^2\,v\left(\frac{0}{N},p,\delta\right)\right]\leq 1$$

This can be further simplified using that, by definition of the average per stage continuation value, $v\left(\frac{0}{N}, p, \delta\right) = 1$.

Using this, and re-arranging.

$$g - \delta g - \delta + \cancel{1} + \left(\frac{N-1}{N}\right)\delta \cdot \left[p^2 v\left(\frac{2}{N}, p, \delta\right) + 2p\left(1-p\right) v\left(\frac{1}{N}, p, \delta\right) + \left(1-p\right)^2\right] \leq \cancel{1}$$

$$\rightarrow \left(\frac{N-1}{N}\right)\delta \cdot \left[p^2 v\left(\frac{2}{N}, p, \delta\right) + 2p\left(1-p\right) v\left(\frac{1}{N}, p, \delta\right) + \left(1-p\right)^2\right] \leq \delta\left(1+g\right) - g$$

We notice that the LHS of the last expression in always greater or equal than zero. This implies that the minimum value of δ that meets this condition is equal or larger than the value of δ of the case every agent was playing a simple trigger strategy against deviations.

Also, lets notice that the Phase I condition is easier to meet as the value of δ grows larger.

Appendix B

Derivation of Incentive Compatibility Condition in Phase II

When an agent that belongs to a set of agents playing according to Phase II is matched with an agent that belongs to a different set of players playing according to Phase I, the following incentive compatibility condition must be met.

$$
\begin{aligned}
(1+g)+\left(\frac{1}{N}\right)\frac{\delta}{(1-\delta)}\cdot 1+\frac{\delta}{(1-\delta)}\left(\frac{N-1}{N}\right)\mathbb{E}_{\mathrm{d}}\left[pv\left(\frac{\mathrm{d}+1}{N},p,\delta\right)+(1-p)\,v\left(\frac{\mathrm{d}}{N},p,\delta\right)\right] &\geq 1\\
+\frac{\delta}{(1-\delta)}\cdot 0+\frac{\delta}{(1-\delta)}\left(\frac{N-1}{N}\right)\mathbb{E}_{\mathrm{d}}\left[pv\left(\frac{\mathrm{d}+1}{N},p,\delta\right)+(1-p)\,v\left(\frac{\mathrm{d}}{N},p,\delta\right)\right]&
\end{aligned}
$$

The LHS of this expression is the sum of the stage payment for playing D and a continuation value in which there is possibly one extra set of players (denoted by the variable d, over which the expectation is defined) playing according to Phase II in the next stage of the game. As there are an arbitrarily big number of players of the same group of players facing a member a member of the same group in any given stage of the game, the LHS is always bigger than the RHS given by playing C plus the same continuation value.

$$(1+g)+\left(\frac{1}{N}\right)\frac{\delta}{(1-\delta)}\geq 1$$

$$\Longleftrightarrow g+\left(\frac{1}{N}\right)\frac{\delta}{(1-\delta)}\geq 0$$

At last, it is easy to see that in the other case (this is, when a player in Phase II is matched with a player of a set also in Phase II) the incentive compatibility condition is always met.

$$-l + \frac{\delta}{(1-\delta)}\left(\frac{N-1}{N}\right)\mathbb{E}_{\mathrm{d}}\left[v\left(\frac{\mathrm{d}}{N}, p, \delta\right)\right] \geq 0 + \frac{\delta}{(1-\delta)}\left(\frac{N-1}{N}\right)\mathbb{E}_{\mathrm{d}}\left[v\left(\frac{\mathrm{d}}{N}, p, \delta\right)\right] + \frac{\delta}{(1-\delta)}\left(\frac{1}{N}\right)$$

Appendix C

Proof of Lemma 1

Following Ellison (1994), in order to prove the convexity of the continuation value, it has to be proven that for any number of sets of agents d that have transitioned to Phase II in a stage t, the following expression is always fulfilled.

$$v\left(\frac{\mathrm{d}}{N}, p, \delta\right) - v\left(\frac{\mathrm{d}+1}{N}, p, \delta\right) \geq v\left(\frac{\mathrm{d}+s}{N}, p, \delta\right) - v\left(\frac{(\mathrm{d}+1)+s}{N}, p, \delta\right)$$

The continuation value for a player j equals the expected value, over a function γ whose outcome is a pairing of every agent in every stage of the game, of a function $f(\cdot)$ whose outcome is the average continuation value per stage for an specific realization of γ.

$$v\left(\frac{\mathrm{d}}{N}, p, \delta\right) = \mathbb{E}_{\gamma}\left[f\left(\frac{\mathrm{d}}{N}, p, \delta, \gamma\right)\right]$$

It is only necessary to prove that for any $s \in \mathbb{N}$ and for any realization of the pairing function γ, the next condition is always satisfied for player j.

$$\mathbb{E}_{\gamma}\left[f\left(\frac{\mathrm{d}}{N}, p, \delta, \gamma\right) - f\left(\frac{\mathrm{d}+1}{N}, p, \delta\right)\right] \geq \mathbb{E}_{\gamma}\left[f\left(\frac{\mathrm{d}+s}{N}, p, \delta\right) - f\left(\frac{(\mathrm{d}+1)+s}{N}, p, \delta\right)\right]$$

Notice that because of the structure of the game, the histories for any player and for any realization of the matching function γ are going to be exactly the same for any $\mathrm{d} > 0$ (or when there already was a first deviation in the game). So the convexity is only strict for $\mathrm{d} = 0$ and with equality for any other $\mathrm{d} \geq 1$.

To prove this, we need to define some sets of groups of players (subsets over the set of groups of players, indexed by $\mathrm{i} \in \{1, \ldots, N\}$). First, we define the set of groups of players that are playing according to Phase II in the stage t of the game if a group M transitions to Phase II in stage 0.

$$O_2(0,\gamma) = \{M\}$$
$$O_2(t,\gamma) = \left\{\mathrm{i} : RD^{\mathrm{i}}_{t-1} \geq p \wedge O_2(t-1,\gamma) \neq \emptyset\right\}$$

In the expression above, $O_2(t,\gamma)$ contains all the sets of players that are playing according to Phase II in t, given that in stage 0 the group M was playing according to Phase II. RD^{i}_t is the realization of the random device of group i in stage t.

We also define $O_1(0,\gamma)$ as the sets of players that in the case described above are playing according to Phase I in stage t. Of course, $O_1(t,\gamma) = O_2^C(t,\gamma)$, $\forall t, \gamma$. Lastly, we define the subsets $S_1(j,\gamma)$ as the list of stages t in which the group of player j plays according to Phase I under a realization γ of the matching function, and $S_2(j,\gamma)$ as its complement.

The effect that one group of players transitioning to Phase II has in the average per stage continuation value of player j has to come from the fact that at least one of his opponents during the game that would have been in Phase I if the marginal group had not entered Phase II, is now going to be playing according to Phase II when encountered by j, and by the fact that during the "distortion" that the deviation of the first group from Phase I causes, the group player j belongs to is going to be switching back and for, according to the realizations of its random device, between Phase I and Phase II.

These effects are reflected in the next equation (where $m_j(t,\gamma)$, lists the opponents in each stage t for the agent j in a given realization of γ). For clarity, please note that because of the structure of the game, the below term will be different from zero only when d is equal to zero and a group M transitions from Phase I to Phase II, starting the contagion.

$$\begin{aligned}
f\left(\frac{\mathrm{d}}{N}, p, \delta, \gamma\right) - f\left(\frac{\mathrm{d}+1}{N}, p, \delta, \gamma\right) &= \sum_{t=0}^{\infty}(1-\delta)\delta^t\left(1-(1-p)^N\right)^t\left[(g)\Phi\left(m_j(t,\gamma) \in O_1(t,\gamma) \wedge t \in S_2\right)\right] \\
&- \sum_{t=0}^{\infty}(1-\delta)\delta^t\left(1-(1-p)^N\right)^t\left[(l+1)\Phi\left(m_j(t,\gamma) \in O_2(t,\gamma) \wedge t \in S_1\right)\right] \\
&+ \sum_{t=0}^{\infty}(1-\delta)\delta^t\left(1-(1-p)^N\right)^t\left[(0)\Phi\left(m_j(t,\gamma) \in O_1(t,\gamma) \wedge t \in S_1\right)\right] \\
&- \sum_{t=0}^{\infty}(1-\delta)\delta^t\left(1-(1-p)^N\right)^t\left[(1)\Phi\left(m_j(t,\gamma) \in O_2(t,\gamma) \wedge t \in S_2\right)\right]
\end{aligned}$$

Where $\Phi(\cdot)$ is an indicator function (has a value of 1 if the condition is met and 0 otherwise). Also, in the last expression the term $\left(\left(1-(1-p)^N\right)^t\right)$ is the probability that the "distortion" caused by the first agent going from Phase I to Phase II lasts until the t stage of the game. The payments in each of the terms above make reference to the difference between the payment of the relevant combination in agents actions and the payment they would received when every group in the game in playing using Phase I. We can use the binomial theorem to re-write $\left(1-(1-p)^N\right)^t$.

$$\left(1-(1-p)^N\right)^t = 1 - \left(\sum_{q=0}^{N}\binom{N}{q}1^{N-q}(-p)^q\right) = \cancel{1} - \left(\cancel{1} + \sum_{q=1}^{N}\binom{N}{q}(-p)^q\right)$$
$$= -\left(p\sum_{q=1}^{N}\binom{N}{q}(-(-p)^{q-1})\right) = \left(p\sum_{q=1}^{N}\binom{N}{q}(-p)^{q-1}\right) = p\cdot R(p,N)$$

Where $R(p,N) = \sum_{q=1}^{N}\binom{N}{q}(-p)^{q-1}$. Taking this, and simplifying the above expression when possible, we can obtain the following.

$$f\left(\frac{\mathrm{d}}{N},p,\delta,\gamma\right) - f\left(\frac{\mathrm{d}+1}{N},p,\delta,\gamma\right) = \sum_{t=0}^{\infty}(1-\delta)\delta^t p^t R^t(p,N)\left[(g)\Phi\left(m_j(t,\gamma)\in O_1(t,\gamma)\wedge t\in S_2\right)\right]$$
$$-\sum_{t=0}^{\infty}(1-\delta)\delta^t p^t R^t(p,N)\left[(l+1)\Phi\left(m_j(t,\gamma)\in O_2(t,\gamma)\wedge t\in S_1\right)\right]$$
$$-\sum_{t=0}^{\infty}(1-\delta)\delta^t p^t R^t(p,N)\left[(1)\Phi\left(m_j(t,\gamma)\in O_2(t,\gamma)\wedge t\in S_2\right)\right]$$

Resuming what we have learned so far, for the continuation value to be convex it suffices that for any realization of the function γ the following condition is met.

$$f\left(\frac{0}{N},p,\delta,\gamma\right) - f\left(\frac{1}{N},p,\delta,\gamma\right) = \sum_{t=0}^{\infty}(1-\delta)\delta^t p^t R^t(p,N)\left[(g)\Phi\left(m_j(t,\gamma)\in O_1(t,\gamma)\wedge t\in S_2\right)\right]$$
$$-\sum_{t=0}^{\infty}(1-\delta)\delta^t p^t R^t(p,N)\left[(l+1)\Phi\left(m_j(t,\gamma)\in O_2(t,\gamma)\wedge t\in S_1\right)\right]$$
$$-\sum_{t=0}^{\infty}(1-\delta)\delta^t p^t R^t(p,N)\left[(1)\Phi\left(m_j(t,\gamma)\in O_2(t,\gamma)\wedge t\in S_2\right)\right] > 0$$

Lastly, notice that strict convexity for every of d could be easily achieved if the value of the probability p of playing Phase II were made an increasing function on the value of d observed in the previous stage ($p_t = p_t(\mathrm{d}_{t-1})$).

Appendix D

Proof of Lemma 2

We have assumed that the payment of the group leaders are proportional to the average stage payments of the members of their respective groups. In case group leader from group I instructs the group to play according to Phase I (PI) in stage t, action that is equivalent to the communication action: $A_t^{I,j} = C$, $\forall j \neq I$, the average discounted payment is given by:

$$P^I(PI, \tau, x) = \mathbb{E}_{\mathrm{d}_t}\left[\left(\frac{\mathrm{d}_t}{M}\right) \cdot (-l) + \left(\frac{M - \mathrm{d}_t}{M}\right) \cdot (1) + V_{\mathrm{d}_{t+1}}(PI, \tau, x) \mid \mathrm{d}_{t-1} = x\right] \tag{D.1}$$

In the last equation, $\left(\frac{M-\mathrm{d}_t}{M}\right)$ is the proportion of groups of agents following Phase I over the total number of groups of players M, and $V_{\mathrm{d}_{t+1}}(PI, \tau, x)$ is the discounted average continuation value from stage $(t+1)$ on, given that in stage t the group I plays according to Phase I, and every group leader is using communication strategy τ in the game from the next stage on. The value of d_t in the equation is the expected value since in stage (t-1) the number of groups following Phase II (d_{t-1}) was equal to x, a known value before the group leader announces what action the group will play in stage t.

On the other hand, if the group leader instructs the group to play according to Phase II (PII), action corresponding to the communication action $A_t^{I,j} = D$, $\forall j \neq I$, the average payment corresponds to the following equation.

$$P^I(PII, \tau, x) = \mathbb{E}_{\mathrm{d}_t}\left[\left(\frac{\mathrm{d}_t}{M}\right) \cdot (0) + \left(\frac{M - \mathrm{d}_t}{M}\right) \cdot (1+g) + V_{\mathrm{d}_{t+1}}(PII, \tau, x) \mid \mathrm{d}_{t-1} = x\right] \tag{D.2}$$

Using this expressions, we can found the condition that must be fulfilled so that the group leader prefers to declare Phase II to the group instead of Phase I.

$$\mathbb{E}_{\mathrm{d}_t}\left[\left(\frac{\mathrm{d}_t}{M}\right)\cdot(-l)+\left(\frac{M-\mathrm{d}_t}{M}\right)\cdot(1)+V_{\mathrm{d}_{t+1}}(PI,\tau,x)\mid \mathrm{d}_{t-1}=x\right]\leq$$
$$\mathbb{E}_{\mathrm{d}_t}\left[\left(\frac{\mathrm{d}_t}{M}\right)\cdot(0)+\left(\frac{M-\mathrm{d}_t}{M}\right)\cdot(1+g)+V_{\mathrm{d}_{t+1}}(PII,\tau,x)\mid \mathrm{d}_{t-1}=x\right]$$

Re-arranging this expression.

$$\mathbb{E}_{\mathrm{d}_t}\left[-\left(\frac{\mathrm{d}_t}{M}\right)\cdot(l)+\left(\frac{M-\mathrm{d}_t}{M}\right)\cdot(1-(1+g))+V_{\mathrm{d}_{t+1}}(PI)-V_{\mathrm{d}_{t+1}}(PII,x)\mid \mathrm{d}_{t-1}=x\right]\leq 0$$
$$\iff \mathbb{E}_{\mathrm{d}_t}\left[-\left(\left(\frac{\mathrm{d}_t}{M}\right)\cdot(l)+\left(\frac{M-\mathrm{d}_t}{M}\right)\cdot(g)\right)+V_{\mathrm{d}_{t+1}}(PI,\tau,x)-V_{\mathrm{d}_{t+1}}(PII,\tau,x)\mid \mathrm{d}_{t-1}=x\right]\leq 0$$

From where, finally, we obtain.

$$\mathbb{E}_{\mathrm{d}_t}\left[V_{\mathrm{d}_{t+1}}(PI,\tau,x)-V_{\mathrm{d}_{t+1}}(PII,\tau,x)\mid \mathrm{d}_{t-1}=x\right]\leq \mathbb{E}_{\mathrm{d}_t}\left[\left(\frac{\mathrm{d}_t}{M}\right)\cdot(l)+\left(\frac{M-\mathrm{d}_t}{M}\right)\cdot(g)\mid \mathrm{d}_{t-1}=x\right] \tag{D.3}$$

The RHS of this last expression is always greater than zero (since g,$l>0$). Now, assume group leaders are using communication strategy $\hat{\tau}$ (meaning that they always play according to the realization of their random device)) from stage $(t+1)$ on. Then the last equation becomes:

$$\mathbb{E}_{\mathrm{d}_t}\left[V_{\mathrm{d}_{t+1}}(PI,\hat{\tau},x)-V_{\mathrm{d}_{t+1}}(PII,\hat{\tau},x)\mid \mathrm{d}_{t-1}=x\right]\leq \mathbb{E}_{\mathrm{d}_t}\left[\left(\frac{\mathrm{d}_t}{M}\right)\cdot(l)+\left(\frac{M-\mathrm{d}_t}{M}\right)\cdot(g)\mid \mathrm{d}_{t-1}=x\right] \tag{D.4}$$

Notice that $\mathbb{E}\left[\mathrm{d}_t\right]$ is independent in this case of the value of d_{t-1} if $\mathrm{d}_{t-1}\geq 1$. The reason for this, is that no matter *how many* of the sets of players are in Phase II in stage $(t-1)$, if there is at least one group playing according to that phase, then all the others are going to draw a realization of its random device in order to see if they have to transition themselves to Phase II.

A direct consequence of this, is that if $x\geq 1$, the particular value of x is non-informative about the value of d_t, meaning that a leader cannot make any *new* inferences about the likelihood of any particular value of d_t.

The latter is very important, because it guarantees that no matter the value of d_{t-1}, the incentives to cheat are always going to be in the same direction at least $\forall \mathrm{d}_{t-1}\geq 1$.

That is, after a first deviation in the game occurs, if group leaders are using an honest communication strategy, they will always want to cheat either by falsely declaring Phase I or by falsely declaring Phase II.

Finally, consider group leaders are using a communication strategy $\tilde{\tau}$, optimal against the review strategy σ defined in *Theorem* 4. We want to argue that in this case group leaders would still always want to deviate either by ordering Phase I or Phase II, no matter the value of d_{t-1}.

A simple way of proving this comes from the fact that if the new communication strategy $\tilde{\tau}$ does not generate a correlation between the number of groups in Phase II in the last stage (d_{t-1}) and the number of groups that will be in Phase II in the following stages of the game, then the change in the symmetric communication strategy is irrelevant. On the other hand, if $\tilde{\tau}$ does generate correlation notice that for any given group leader, because of the stage game being a Prisoner's Dilemma, if the value of d_{t-1} makes very likely that in the next stage there are going to be a big fraction of the groups in Phase II, then the optimal deviation from $\hat{\tau}$ is to falsely declare Phase II, while if it is the opposite case, then the optimal deviation corresponds exactly to the *same lie*. Therefore, group leaders want to deviate in the same way, no matter the value of d_{t-1} they saw, even though they can now somehow predict the behaviour of the rest of the groups in future stages of the game.