Università
Ca'Foscari
Venezia

Corso di Dottorato di ricerca
in Economia
Ciclo XXXI

Tesi di Ricerca

Essays on Coarse Cognition, Competition, and Cooperation

SSD: SECS-S/06

Coordinatore del Dottorato
ch. prof. Giacomo Pasini

Supervisore
ch. prof. Marco LiCalzi

Secondo Supervisore
ch. prof. Giacomo Pasini

Dottorando
Matteo Aggio
Matricola 828304

CA' FOSCARI UNIVERSITY OF VENICE

Abstract

Department of Economics

PhD in Economics

Essays on Coarse Cognition, Competition, and Cooperation

by Matteo Aggio

This dissertation is divided in three self-contained and single-authored chapters. We analyze, in three different theoretical models, the role of coarse cognition, also defined as categorical rationality, and its consequences on agents' behavior and incentives. In the first and the second chapter, we analyze a model of competition between cognitively limited agents (or firms). Specifically, in the first chapter we investigate the consequences of competition between agents that have different abilities to recognize the states of the world. In the second one, we provide necessary conditions for agents to cooperate, either tacitly or explicitly. The third chapter presents a cooperation model where agents are compelled to redefine their communication in order to meet the cognitive limitations of the partner. As a general result, we find that coarse cognition has deep consequences on agents' ability to understand the situations they face, affecting their ability to implement precise strategies, that are optimal given agents' information but suboptimal with respect to the true state of the world.

Introduction

In economic theory, the assumption of perfect rationality have long been deemed as too strict and unrealistic, and recently literature has, both theoretically and experimentally, tried to move on from this assumption. The aim of this dissertation is to propose economic applications of what we call *coarse cognition*, i.e. the incapacity to perceive reality precisely as it is. Due to their cognitive limitations, agents categorize the information they receive, bundling together similar states. In three different chapters, we present three theoretical models in which we analyze agents' strategic behavior when they do not fully recognize or understand the world where they make their decisions. In the first and in the second chapter, we evaluate the role of cognitive limitations in a competition setting, while in the third chapter we consider a cooperation game.

Chapther 1 investigates the impact of categorical rationality on strategic choices. In particular, in the framework of a Cournot duopoly, we investigate the implications of agents' tendency to bundle together similar situations (as in [Mengel, 2012], [Heller and Winter, 2016], and [LiCalzi and Mühlenbernd, 2019]). We find that the best joint-performing dyads are composed of agents with individual categories that cover different intervals of the market demand, and that they may outperform agents with the ability to choose their own categories. Further, we analyze communication, finding that only unilateral communication can make the arising cognitive advantage greater than the competition disadvantage.

Chapther 2 introduces agents' coarse cognition in an industrial organization setting, where we ivestigate firms' incentives to start and maintain tacit or explicit collusion in the long run. Our analysis is applied to a repeated Cournot game. We assume that firms, as in the first chapter, have cognitive limitations that impair their perception of the market demand. While the literature has extensively investigated firms' incentives to form cartels (see [Levenstein and Suslow, 2006] for an empirical survey, and [Harrington, 2017] and [Garrod and Olczak, 2018] for more recent theoretical findings), not much is known on the role of the perception of the market demand and how it affects firms' capacity to collude, either tacitly or explicitly. Consistently with the emergent literature, we find that tacit collusion is, in the long run, easier to sustain than explicit collusion, despite being, under certain circumstances, less profitable (see [Fonseca and Normann, 2012] and [Waichman et al., 2014] for similar experimental findings). Finally, we perform a welfare analysis, and we find the conditions under which the Competition Authority could mitigate the negative effects of collusion.

Chapther 3 provides an explanation for performance heterogeneity among teams with ex ante similar initial conditions. We study a model of team-working where two agents form a dyad and decide whether to launch or not a two-dimensional project, which is potentially profitable. We investigate two sources of inefficiencies. On the one hand, each agent is expert in a specific fields, and has cognitive limitations regarding the other field: as they are aware of each others' limitations, agents are compelled to redefine their communication at a coarser level (as in [Wernerfelt, 2004]). We find that categorical communication always decreases expected profits compared to perfect communication, and that there exist ways in which these inevitable inefficiencies can be mitigated. On the other hand, we study the effects of biased beliefs of team members: agents may overestimate or underestimate the profitability of the project, and hence increase the probability of incurring wrong choices. We find that these biases not only decrease expected profits, but may also make communication between team members detrimental.

Acknowledgements

The journey I am about to finish left me as a more mature person (but not necessarily wiser) with respect to the one that started it, almost four years ago. The course has had its ups and downs, with its moments of joy and difficulty, nonetheless every occasion has been an important step to grow as a researcher and learn new things. Hence, I heartily thank all the people that shared at least one of these meaningful moments with me.

I warmly thank my supervisor Marco LiCalzi for the fundamental guidance held during these years. His mentoring has led me to improve and raise my goals as a researcher, and I will carry his advice with me for the years to come.

I deeply thank my second supervisor Giacomo Pasini for his support in all the aspects related to this long journey. He helped me with his constant presence and I will always be grateful to him.

I am extremely grateful to the members of the committee, Valeria Maggian, Ennio Bilanici, and Stefano Galavotti for reading my work (twice) and for exerting sincere effort and commitment in order to improve it. Their precious suggestions helped me read between the lines of my work, inspiring improvements, ideas and possible extensions. Their work has been central for the final development of this thesis.

I would like to thank the whole Department of Economics for supporting me from the very first day, and the MIT Sloan School of Management for hosting me during my visiting period.

This PhD has been an occasion for getting to know some wonderful persons. Firstly, and with all my hearth, I would like to thank Caterina, the person to whom I owe the most. Without any doubt, meeting you has been the best gift this journey has brought me. I would like to give a big hug to my brother Sebastiano, with whom I shared this PhD from the beginning to the end, always together. Your friendship has an immeasurable value, and it has supported me when I needed the most. I thank my colleagues, who have been a great inspiration for all this time. The days spent altogether, trying to improve our research and to figure out our goals, represents an invaluable memory. From the bottom of my hearth, I would like to thank Enrico for letting me be his younger brother during my visiting period at MIT.

I thank my friends Andrea, Tommaso, Claudio, Nicola, Edoardo, and Daniele for always being such a great group of friends. The years go by, we live our lives (distant, by now) but nothing changes. And this is due to the extremely strong ties that link us.

I dedicate this work to my family. I know sustaining me has been hard work, so I praise your patience and thank you for your constant support. All of this would not have been possible without you.

Contents

List of Figures

Chapter 1

Categorical Rationality and Competitive Interaction

Abstract

This chapter builds on a recent literature that investigates agents with categorical rationality: agents that maximize utilities given their categorized information. In our setting, agents categorize bundling together different states of the world. They lose informational precision, but they reduce the number of strategic decisions to make. We analyze a one-shot Cournot competition where two agents categorize the market size into two exogenous and fixed categories: they cannot distinguish values of the market size that belong to the same category. We find that the best joint-performing dyads are composed of agents with individual categories that cover different intervals of the market size, and that they may outperform agents with the ability to choose their categories. We then examine communication and the trade-off it generates between a cognitive advantage and a competition disadvantage. The latter usually overcomes the former, unless communication is unilateral and other certain conditions hold. This chapter contributes to the literature that investigates the implications of agents with categorical rationality. Overall, our results contribute to shed light on how the coordination (or the miscoordination) of categories can affect agents' collective performance.

Keywords: Categorical rationality, Competition, Categorization, Communication.

JEL Classification Numbers: D83, D91.

> *Reflecting the world, we organize it in entities: we analyze the world gathering and breaking up a continuum of almost uniform and stable processes, to better deal with them. [...] It is the structure of our nervous system that works in this way.*[1] *[Rovelli, 2017]: 149.*

1.1 Introduction

When they perceive a new phenomenon, people tend to relate it to phenomena observed in the past that have analogous characteristics. This mental operation is called categorization and is one of the most basic processes that regard cognition. Stated differently, categorization implies classifying objects and ideas into disjoint classes called categories, that help the human brain deal with vast amounts of information: accommodating individuals' cognitive limitations, categories reduce the costs associated with gathering new information, storing and registering it ([Baumol and Quandt, 1964]).

Given its relevance, the study of categorization spans several disciplines, ranging from linguistics to neuroscience, psychology, and economics. In economics, categorization has been investigated both in decision and in game theory, and more specifically in the literature regarding bounded rationality. Besides the importance of the topic, many salient aspects still need to be examined. Up to now, two main strands of literature have emerged on this subject. The first focuses on the evolutionary formation of categories (see for instance [Mengel, 2012], [Heller and Winter, 2016], [LiCalzi and Mühlenbernd, 2019]). The second one investigates the implications and empirical evidence of categorization ([Samuelson, 2001], [Halevy et al., 2012], [Gibbons and Henderson, 2013] among others). Our paper contributes to this second branch of literature, that offers a variety of theoretical approaches and results.

In our paper, we aim at exploring how agents with categorical rationality (i.e. agents that maximize their utility conditional on their categories) are reciprocally influenced by their different categorizations. In particular, we study the case of two agents involved in a competitive game with continuous strategies (assuming a Cournot competition). We analyze individuals who coarsely categorize the market demand, and have exogenous and fixed categories. *Exogenous* means that agents are endowed with specific categories, *fixed* because categories cannot be modified.[2]

More specifically, in our model we assume that the true market size of the Cournot competition is not directly observed by the agents, who can only discriminate between two categories. In this case, a category is an interval of possible values for the market size. Further, agents need not share their categorization with the opponent: they learn them before the game starts.

Our analysis considers four different steps. First, we compare pairs of agents with different categorizations, and we compute which of these dyads makes higher profits on average. We find that the best performing dyads are composed of agents with individual categories that allow them to specialize in different intervals of the market size. Second, we proceed by comparing agents who have two fixed categories with agents that have the ability to set their own two categories. Under certain conditions, the second type

[1] Author's translation from Italian. The actual reference is: "Nel riflettere il mondo, lo organizziamo in enti: pensiamo il mondo raggruppando e spezzettando al meglio un continuo di processi piú o meno uniformi e stabili, per meglio interagire con essi. [...] É la struttura del nostro sistema nervoso che funziona in questo modo".

[2] As an explanatory example, think about driving speed in highways. People roughly discriminate between two cases: low and high driving speed, where the boundary is set to 130 km/h. How we categorize driving speed (low below 130 km/h, high above it) is exogenous (the government set the rule) and fixed (the drivers cannot modify it).

of agents may perform worse than agents with exogenous and fixed categories. In fact, the stage-game of deciding categories resembles a prisoners' dilemma: agents make a decision that is individually optimal but not socially efficient. Third, we analyze whether truthful communication can help agents and lead them to higher payoffs. Because of the competitive nature of the game, we find that communication generates a trade-off between an improved perception of the game (a cognitive advantage given by the fact they have more information and they may refine the understanding of the market size) and an increase in the strength of the competition within each situation (competition disadvantage). Results show that, in truthful bilateral communication, the competition disadvantage always overcomes the cognitive advantage; if they communicate bilaterally, agents may end up being worse off. Thus, more information is not always better. Lastly, analyzing the intermediate case of truthful unilateral communication, we find that it can yield to a Pareto improvement. Under specific conditions and assuming transferable utility, having only one agent that communicates is beneficial for both players.

Overall, our results contribute to shed light on how the coordination (or the mis-coordination) of categories affects agents' joint performance. This can be relevant in industrial organization, and especially in the study of duopolies: decision makers could increase their profits (and decrease consumers' surplus) depending on how they perceive the market demand.

This chapter is organized as follows. The rest of this section assesses the related literature. Section 1.2 presents the model and the assumptions. Section 1.3 introduces binary categorizations, and derives the first two results: we introduce agents' equilibrium strategies and compare expected payoffs for agents with exogenous versus endogenous categories. In Section 1.4 we introduce truthful communication, and we look for conditions under which bilateral or unilateral communication can yield a Pareto improvement. We find such conditions only for the latter. Section 1.5 examines the assumption of truthful communication. We find that agents have incentives to lie and end up avoiding communication at all. Section 1.6 concludes. Appendix A includes the proofs of the propositions stated along the paper. Appendix B encompasses the special case of agents with shared categories.

1.1.1 Related literature

As [Allport, 1954] noted, the process of categorization is inevitable and helps the human mind to think and interpret whatever information is registered. Cognitive scientists tend to agree that categorization is one of the most basic phenomena of cognition, and it represents a necessary mental step that individuals make in order to organize their knowledge (see [Cohen and Lefebvre, 2005] for an extensive inquiry). Individuals need easily accessible tools to understand and use a new piece of information. Categories represent a very useful shortcut: they accommodate individuals' cognitive limitations and let them classify the upcoming information depending on past experiences.[3]

As noted, two main strands of literature have emerged about categorization. Some articles have analyzed the implications and empirical evidence. For instance, [Samuelson, 2001] analyzes agents that play different types of games (ultimatum, bargaining, and tournament) and finds that the first two can be grouped together in order to save on complexity costs. A model of categorical rationality is used by [Gibbons and Henderson, 2013] to find that when players share categories in a one-shot interaction they may

[3]Some kinds of categorizations even seem to be independent of sensory experience: for instance, the structural rules behind Chomsky's Universal Grammar, or basic object categorization, as stated in [Hoff and Stiglitz, 2016]: 26.

end up better (or worse) off than when they perfectly discriminate the space of games. Using categorical rationality, [Mullainathan et al., 2008] explains how a persuader (e.g. a product advertiser) can take advantage of the people she tries to persuade: she can exploit categorical thinking of target agents and send uninformative messages that shift the attention of receivers to other topics, much more apt to influence receivers' behavior. [Halevy et al., 2012] find empirical results that support the idea that individuals categorize strategic interactions: a natural tendency to simplify induces people to perceive only a limited number of archetypal situations (only four). An optimal degree of refinement in categorical thinking is assessed by [Mohlin, 2014], who finds that optimal number and shape of categories are endogenous to the model and solve the trade-off between category homogeneity and the ability to make better predictions.

More recent articles appraise *how* categories are formed: the underlying idea is that categories are molded by evolutionary processes. For instance, in [Mengel, 2012], agents must learn which games to discriminate, and they need not always end up learning the same partition. [Mengel, 2012b] tests the evolutionary fitness of different partitions (i.e. categorizations), finding conditions under which coarse partitions have higher evolutionary fitness than finer partitions. [Heller and Winter, 2016] investigate the question of rule rationality: the properties of the strategic environment determine which kind of rule of thumb (or mode of behavior) agents decide to apply. In this paper, the authors prove that bundling together many decision situation can be beneficial, net of smaller cognitive costs and fewer informational requirements. [LiCalzi and Mühlenbernd, 2019] depart from these ideas assuming agents cannot choose their categorizations, and their main result is that, in the long run, agents end up sharing their categorization abilities. The view that categorization is the result of an evolutionary process is coherent with [Bednar and Page, 2007], for which ensembles of games can result in "aggregate behavior that can be defined as cultural", and with [Hoff and Stiglitz, 2010], who argue that social beliefs systems are rigid and difficult to modify.

There is a wide literature on the role of society in shaping categorical rationality. As in [Denzau and North, 1994], individuals with common cultural background will reasonably share similar ideologies and theories to interpret their environment. [Hoff and Stiglitz, 2016] argue that the set of mental models (tools that individuals use to process information and conceptualize) is actually a product of the social context, and individuals draw their mental models from this set: "in this sense, exposure to a given social context shapes *who people are*" ([Hoff and Stiglitz, 2016]: 26). The use of specific categories may depend on the society in which an individual operates: as the development of *social norms* ("behavior of a group if the behavior differs from that of other groups in similar environments", [Benhabib et al., 2011]: 32), the evolution of categories is a lengthy process whose outcome may be difficult to modify or eradicate. [Fryer and Jackson, 2008] provide a link between categorization and social interactions: for instance, minorities are categorized more coarsely than the majority.

1.2 The model

Consider two agents, labeled 1 (male) and 2 (female), engage in a Cournot competition game. Each agent $i = 1, 2$ simultaneously chooses a production quantity x_i facing a null marginal cost; $x_i + x_j$ is the aggregate output quantity, and we assume agents face an inverse linear demand. The market demand is

$$d = a - (x_i + x_j)$$

where a is the value of the market size. Payoffs depend on both agents' production decisions and the market size:

$$\pi_i = \max\,\{x_i d, 0\} \tag{1.1}$$

We assume that the market size is a random variable A. We let a be the true state, i.e. the realization of the market size A.[4] Let $R_A = (0,1)$ be the support of A, i.e. the set of all possible values of A. Note that the support of A is a totally ordered set. Moreover, for simplicity, we assume that A is uniformly distributed. All of the above is common knowledge. As an explanatory example, we can think of two duopolists who need to set their production, but do not precisely know how large is the market demand going to be in the following period.[5]

Categorization. Before the game starts, each agent $i = 1,2$, due to its cognitive limitations, categorizes a in two alternative subsets of R_A. A *categorization* C_i over R_A, for the agent $i = 1,2$, is a partition of R_A. Denote the cardinality of C_i by $\#C_i$ (sometimes we refer to $\#C_i$ as the *discrimination level* of agent i). For simplicity, we assume $\#C_i = 2$, for each $i = 1,2$: agents have a *binary categorization.* That is, each agent i partitions the support R_A in two *categories*, which we denote C_i^L and C_i^H (as R_A is totally ordered, categories are intervals).[6] Hence, $C_i = \{C_i^L, C_i^H\}$, where the labels L and H denote *Low* and *High* market sizes, respectively: in fact, without loss of generality, we assume $\sup C_i^L < \inf C_i^H$. Note that by definition of categorization, we have $C_i^L \cap C_i^H = \emptyset$ and $C_i^L \cup C_i^H = R_A$. Moreover, we assume categories are convex intervals.

We assume that agents are *categorically rational*: they are rational, conditional on the information embedded in their categories. Because agents have cognitive limitations, they cannot perfectly discriminate the market size A, and are unable to detect the exact value a. For a given categorization C_i, an agent is able to tell only which category (i.e., which region of the support R_A) contains a.

This implies that the agents have only partial information about the game they are about to play. Thus, cognitive limitations affect agents in two ways: (i) directly, by influencing their expectation about the market size, and (ii) indirectly, through their expected market demand and profits.

At the beginning of the game, nature chooses the market size a from a given distribution. Due to its cognitive limitations, each agent recognizes only a category (i.e. a block) containing a. *Ex ante*, an agent i has only two complementary categories C_i^L and C_i^H. *Ex post*, we let S_i^a denote the category that includes a. We call S_i^a the *situation* perceived by agent i once a is realized (the superscript a reminds us that S_i^a is the category that includes a). For instance, if $a \in C_i^L$, then $S_i^a = S_i^L = C_i^L$. In general, we have that

$$S_i^a = \begin{cases} S_i^L \text{ if } a \in C_i^L \\ S_i^H \text{ if } a \in C_i^H \end{cases}$$

Categories are given *ex ante*, before the market size is realized; *situations* are perceived *ex post*, after the market size is realized.

[4] Throughout the dissertation, we denote random variables by upper-case latin letters and their realizations by lower-case latin letters.

[5] For instance, two rental cars who have to forecast the demand for cars each month: they cannot figure out the exact number of cars demanded because of several unpredictable conditions (e.g. weather conditions or strikes in the public transportation service).

[6] This can be considered a particular case of [Mengel, 2012], where the marginal cost $\xi_j = 0$ for the categories $j = 1,2$, and $\xi_j = \infty$, for all $j \geq 3$.

Example 1.2.1. *Assume* $C_i^L = (0, 0.4]$ *and* $C_i^H = (0.4, 1)$ *are the categories through which agent* i *discriminates the support of the market size* R_A. *As the game starts, let* $a = 0.41$ *be a realization of* A. *Then, agent* i *perceives situation* S_i^H: *she knows that* $a = 0.41$ *lies in the category* $C_i^H \subset R_A$, *but she does not know its exact value.* △

Thresholds. Each agent has her own categorization of the market size. The mechanism through which each agent $i = 1, 2$ categorizes the market size can be described by a *threshold* parameter $\alpha_i \in [0, 1]$, exogenous and fixed, that separates the two contiguous categories. In particular, categories are $C_i^L = (0, \alpha_i)$ and $C_i^H = (\alpha_i, 1)$. Thus, when $a < \alpha_i$, agent i perceives the low situation S_i^L. When $a > \alpha_i$, she perceives the high situation S_i^H.

For instance, if α_i is very close to 0, then agent i would bundle together all the games where the market size is not negligible. By contrast, when α_i is close to 1, then the agent would bundle together games where the market size is very large. In the limit case where α_i is either 0 or 1, there is a unique category that bundles all possible values. Given the threshold α_i and the situation S_i^a, each agent's expectation about the market size is given by $E_i[A|S_i^a]$ and is a function of α_i.

We assume that agents know how the opponent categorizes (and hence her threshold) before the game starts. Moreover, without loss of generality, we assume $\alpha_1 \leqslant \alpha_2$. Although an agents bundles two different realizations of A in the same category, she still knows that her opponent may treat these two realizations differently.

Example 1.2.2. *Assume* $\alpha_1 = 0.4$, $\alpha_2 = 0.6$, *and* $a = 0.41$. *Then, agent 1 perceives the H situation, whereas agent 2 perceives L. We have* $E_1[A|S_1^H] = \frac{\alpha_1+1}{2} = 0.7$, *and* $E_2[A|S_2^L] = \frac{\alpha_2}{2} = 0.3$. △

Communication. After perceiving the situation S_i^a, each agent i can share this piece of knowledge with the opponent by sending a message. Let $M_i = \{M_i^L, M_i^H, \emptyset\}$ be the set of possible messages for each sender $i = 1, 2$, where $M_i = \emptyset$ denotes a void message.[7] A (non-void) message sent by agent i is an element of the set of categories of i; that is, if $M_i \neq \emptyset$, then $M_i \in C_i$. Thus, for a generic message M_i, we have $M_i \in \{C_i \cup \emptyset\}$.

We assume that non-void messages are truthful, and this is common knowledge: messages must contain the true state of the world, that is $M_i = S_i^a$. Hence, we assume that agent i reveals the opponent exactly the situation she has perceived, i.e. the category that contains a. Moreover, agents know that messages are truthful, hence they believe the message they receive.

When agent i receives a message M_j from the opponent j, she gathers additional information about the market size. In fact, as long as she knows the message is truthful, she may be able to reduce the interval in which she believes a is included from S_i^a to $S_i^a \cap M_j$.

We say that agent i *refines* her understanding of the market size if the message allows her to shrink the set of possible values of the market size. This updating refines the understanding of the market size only if the following conditions are satisfied: (i) $M_j \cap S_i^a \neq \emptyset$, and (ii) $M_j \not\supseteq S_i^a$. These conditions guarantee that the combination $M_j \cap S_i^a$ reduces the possible values of A perceived in S_i^a, i.e. $(M_j \cap S_i^a) \subset S_i^a$.

Example 1.2.3. *Let* $\alpha_1 = 0.4$, $\alpha_2 = 0.6$, *and* $a = 0.41$. *When agents truthfully exchange their information, agent 1 sends a message* $M_1 = M_1^H$ *whereas agent 2 sends* $M_2 =$

[7] A void message carries no information at all.

M_2^L. Then, both agents refine their understanding of A: agent 1 from $E_1\Big[A|S_1^H\Big] = 0.7$ to $E_1\Big[A|S_1^H \cap M_2^L\Big] = \frac{\alpha_1+\alpha_2}{2} = 0.5$, agent 2 from $E_2\Big[A|S_2^L\Big] = 0.3$ to $E_2\Big[A|S_2^L \cap M_1^H\Big] = \frac{\alpha_1+\alpha_2}{2} = 0.5$. Now, suppose $a = 0.39$; nothing changes regarding the situation perceived (and the message sent) by agent 2, but agent 1 now perceives S_1^L, expecting A to be $E_1\Big[A|S_1^L\Big] = \frac{\alpha_1}{2} = 0.2$. He sends the message M_1^L, and hence agent 2 updates her expectation $E_2\Big[A|S_2^L \cap M_1^L\Big] = \frac{\alpha_1}{2} = 0.2$. Agent 1 updates his expectation given M_2^L but he does not refine his understanding of A: in fact, $E_1\Big[A|S_1^L\Big] = E_1\Big[A|S_1^L \cap M_2^L\Big] = 0.2$. A similar reasoning applies to agent 2 when $a > 0.6$. △

Truthful messages are an assumption. If we relax this assumption, we let the agents have the possibility to communicate strategically, i.e. convey the message they want regardless of the perceived situation; hence, if we let agents choose their messages in order to maximize their expected payoffs, both agents may have incentives to deceive the opponent and send untruthful messages. When this is the case, the receiver of the message would rationally believe that the message itself may be untruthful. We analyze this possibility later in Section 1.5.

Expected payoffs. When they set their production quantities, agents decide dependently on the market demand they expect. The expected market demand for agent i is

$$E_i\Big[D\Big|S_i^a, M_j\Big] = E_i\Big[A|S_i^a \cap M_j\Big] - \Big(x_i + E_i[x_j|S_i^a, M_j]\Big) \tag{1.2}$$

and depends on the situation perceived by agent i (S_i^a), the message received by the opponent (M_j), the expected quantity set by the opponent ($E_i\Big[x_j|S_i^a, M_j\Big]$), and the own quantity x_i. Thus, the expected payoff function for agent i is

$$E_i\Big[\Pi_i\Big|S_i^a, M_j\Big] = \max\Big\{x_i E_i\Big[D\Big|S_i^a, M_j\Big], 0\Big\} \tag{1.3}$$

The payoff attributed by i to the combination $S_i^a \cap M_j$ of situation S_i^a and message M_j is the expected value across all compatible market sizes. To distinguish it from the *ex post* payoffs in (1.1), we call (1.3) the *ex ante* expected payoff if neither S_i^a nor M_j have yet been observed, or the *interim* expected payoff if only S_i^a or both S_i^a and M_j have been observed.

All agents behave rationally given their beliefs: they set their production quantities after they have maximized the *interim* expected payoffs. Once nature has drawn a and production decisions are made, *ex post* payoffs are realized; then, the true market size is eventually observed.

Timing. The timing of the game is the following:

1. nature draws a, realization of A;

2. each agent $i = 1, 2$ privately observes S_i^a and derives her expectation $E_i\Big[A|S_i^a\Big]$ about the market size;

3. agents may simultaneously exchange messages: when an agent i receives a message M_j, she updates her expectation $E_i\Big[A|S_i^a \cap M_j\Big]$;

4. each agent makes her production decision x_i in order to maximize her *interim* expected payoffs;

5. payoffs are realized.

As a benchmark, we analyze the game under (i) no discrimination, and (ii) perfect discrimination (or perfect categorization). Under no discrimination, agents have only one coarse category (as if $\alpha_i \in \{0,1\}$). Under perfect categorization they can discriminate every value of the market size. Corresponding to the first case is the common *meet* of all the partitions of R_A, ordered by refinement, whereas the latter case corresponds to their *join*.

Benchmark 1: no discrimination. The meet is the coarsest categorization of R_A ($\#C_i = 1$); hence, when agents do not discriminate, they only perceive $S_i^a = R_A$ and infer that $A \sim \mathcal{U}(0,1)$. We call this benchmark *no discrimination.* Agents' expectation about the market size is $E_i[A|S_i^a] = \frac{1}{2}$. Communication is useless, because messages cannot refine agents' understanding. As they both maximize their profits, agents have the following reaction functions:

$$\begin{cases} x_1 = \frac{E_1[A]-x_2}{2} \\ x_2 = \frac{E_2[A]-x_1}{2} \end{cases}$$

Solving the system, we get the optimal production decisions x_i^M (where the superscript M is a reminder for "meet") for both agents: $x_1^M = x_2^M = \frac{1}{6}$. Then, both the *ex ante* the and *interim* expected profit functions are

$$E_i\left[\Pi_i^M\right] = \frac{1}{36}$$

Expected profits are constant because agents forecast the same value for any realization of A. Hence, *ex post* payoffs in (1.1) only depend on a (and they are increasing):

$$\pi_i^M(a) = \max\left\{\frac{3a-1}{18}, 0\right\}$$

Benchmark 2: perfect categorization. The join is the finest partition: both agents can perfectly discriminate every realization of A, as if they had an infinite number of categories. We call this case *perfect categorization.* The reaction functions

$$\begin{cases} x_1 = \frac{a-x_2}{2} \\ x_2 = \frac{a-x_1}{2} \end{cases}$$

are increasing in the realized size a of the market. Solving the system yields the optimal productions: $x_1^J = x_2^J = \frac{a}{3}$. The corresponding profits are:

$$\pi_i^J(a) = \frac{a^2}{9}$$

As actions, profits are increasing in a: the greater the size of the market, the greater the profits, that span from 0 to $\frac{1}{9}$.

1.3 Binary categorization

We analyze the game played by two agents with different binary categorizations, and hence different thresholds. Momentarily, we leave out the possibility that agents exchange (non-void) messages (i.e. we have that $M_i = \emptyset$ for each $i = 1, 2$): this option is analyzed later. We first delineate the equilibrium production functions of agents with binary and fixed categorization, analyzing their expected profits. We then compare the equilibrium payoffs of agents endowed with exogenous and fixed categories with the payoffs of agents with full control on their thresholds, and we find conditions under which the first ones perform (on average) better than the second ones.

Given the assumptions stated in Section 1.2 (A is uniformly distributed and $\alpha_1 \leqslant \alpha_2$), it is possible to represent the agents' categorization of the market size as in Figure 1.1.

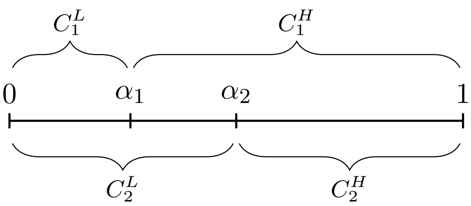


FIGURE 1.1: Example of agents' heterogeneous categorization of R_A.

Given their coarse categorization, agents choose actions for the two situations they can perceive: call x_i^k the quantity produced by agent i when she perceives the situation S_i^k, where $k \in \{L, H\}$. Hence, i decides to produce a low quantity x_i^L after perceiving a low market size and a high quantity x_i^H after perceiving a high market size.

If agents had the same threshold, they would perceive the same situation. When agent have different thresholds, this is no longer true: when they set their production quantities, agents need to take into account the situation potentially perceived by the opponent. In fact, each agent knows that two possible values of a that she cannot discriminate might conversely be perceived differently by the other. For instance, if agent 1 perceives that the market size is large (S_1^H), he knows that agent 2 may perceive either S_2^L or S_2^H, but he cannot tell in advance.

Let $E_i[x_j|S_i^a]$ be the production quantity that i, after perceiving situation S_i^a, expects the opponent j to produce. Clearly, $E_i[x_j|S_i^a]$ is a convex combination of x_j^L and x_j^H:

$$E_i\Big[x_j|S_i^a\Big] = \Pr\Big(S_j^a = S_j^L|S_i^a\Big)x_j^L + \Pr\Big(S_j^a = S_j^H|S_i^a\Big)x_j^H \tag{1.4}$$

Given (1.4), and combining (1.2) and (1.3), after perceiving situation S_i^a the maximization problem for agent i becomes

$$\max_{x_i^k} \; E_i\Big[\Pi_i|S_i^k\Big] = \max\Big\{x_i^k\Big(E_i\Big[A|S_i^k\Big] - (x_i^k + E_i\Big[x_j|S_i^k\Big]\Big)\Big), 0\Big\} \tag{1.5}$$

where $k \in \{L, H\}$.

The maximization of *interim* expected profits in (1.5) yields a system of four equations in four variables (2 situations $\times$ 2 agents): the solutions are the equilibrium production functions. Let x_i^{k*} be the equilibrium production for agent i after having observed situation S_i^k, where $k \in \{L, H\}$: x_i^{k*} only depends on thresholds α_1 and α_2. Substituting the equilibrium productions in the profit function (1.5), we find the equilibrium *interim* expected payoffs $E_i\Big[\Pi_i^*|S_i^a\Big]$; weighting the *interim* expected payoffs by the probability that i observes each situation yields the *ex ante* expected payoffs $E_i\Big[\Pi_i^*(\alpha_1,\alpha_2)\Big]$. Once the market size is realized, agent i earns payoff $\pi_i^*(a,\alpha_1,\alpha_2)$.

Proposition 1.1. *Agent i produces the equilibrium quantity x_i^{k*} when she perceives situation S_i^k, where $i = 1, 2$ and $k \in \{L, H\}$. Ex ante expected payoffs are $E_i\Big[\Pi_i^*(\alpha_1,\alpha_2)\Big]$ and depend on thresholds α_1 and α_2.*

Proof. See Appendix A. □

When agents have different categorization of the market size (and in particular $\alpha_1 < \alpha_2$), each of them has a distinctive competitive advantage: agent 1 perceives a finer L situation than the one perceived by agent 2 ($C_1^L \subseteq C_2^L$); as his category for low market sizes is finer, agent 1 has more precise expectations about the market size than agent 2 when a is sufficiently low. The same happens for agent 2 when the market size realization is large enough. Agents make the most of their discrimination ability: they take more fine-grained actions when they deal with finer situations.[8]

Example 1.3.1. *Suppose agent 1 has a low threshold ($\alpha_1 = 0.2$) and agent 2 a high one ($\alpha_2 = 0.8$). If nature draws $a = 0.1$, they both perceive the L situation. If they could perfectly discriminate across market sizes, the best action would be choosing $x = \frac{a}{3} = 0.03$ (as in Benchmark 2). Instead, agent 2 produces $x_2^{L*} = 0.12$, about four times the optimal production under perfect discrimination. Agent 1, knowing the discrimination ability of agent 2 and her propensity to overproduce, nullifies his production in order not to lose money. An opposite situation occurs if nature draws, for instance, $a = 0.9$. The optimal production decision under perfect discrimination is $x = 0.3$; as both perceive the H situation, agent 1 underproduces setting $x_1^{H*} = 0.21$, whereas agent 2 adjusts her production setting $x_2^{H*} = 0.34$.* △

Best binary categorizations. Both under no discrimination and perfect categorization, *ex post* payoffs only depend on the market size a: in these two benchmarks, given agents play their equilibrium strategies, *ex post* payoffs range from $\pi_i = 0$ when $a = 0$ to $\pi_i = 11.1$ when $a = 1$.[9]

Under binary categorization, realized payoffs $\pi_i(a,\alpha_1,\alpha_2)$ in (1.1) depend not only on a but also on thresholds α_1 and α_2. To find the spectrum of possible payoffs, we identify the minimum and maximum payoffs. Let $\underline{\pi}_i$ and $\bar{\pi}_i$ be, respectively, the global minimum and maximum of $\pi_i(a,\alpha_i,\alpha_j)$ for agent i given a, α_1 and α_2 are in $(0,1)^3$. The minimum realized payoff $\underline{\pi}_i$ is zero when $a = 0$. To find the maximum payoff $\bar{\pi}_i$, we consider the space $[0,1]^2$, where each point represents a pair of thresholds (α_1,α_2), or, similarly, a dyad with thresholds (α_1,α_2). Given that $\pi_i^*(a,\alpha_1,\alpha_2)$ is increasing in a, we hereafter show, in Figure 1.2, the contour plots for *ex ante* expected payoffs $E_1\Big[\Pi_1^*(\alpha_1,\alpha_2)\Big]$ and $E_2\Big[\Pi_2^*(\alpha_1,\alpha_2)\Big]$.

The contour plots suggest that agent i expects the highest possible profit when (i) she has a threshold $\alpha_i = 0.5$, and (ii) opponent j has a *boundary* threshold, i.e. a threshold that coincides with a boundary of R_A. This implies that the best achievable individual

[8] For instance, the cross-race effects deal with a related question: ethnic groups recognize members of their own better than members of other ethnic groups. See [Fryer and Jackson, 2008] and [Tanaka and Taylor, 1991] for a more extensive inquiry.

[9] Hereafter we multiply the realized payoffs by 100.

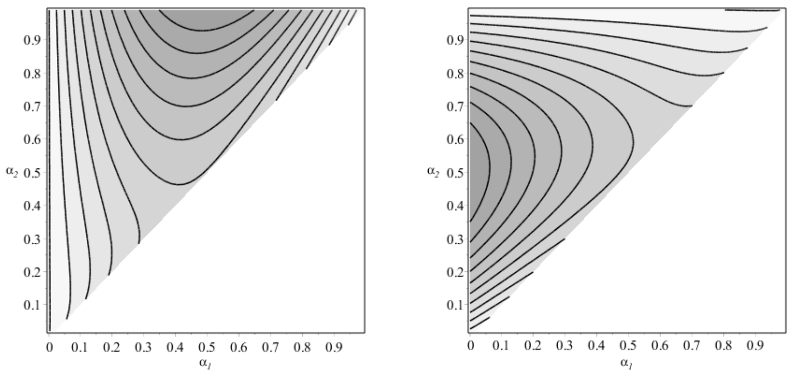


FIGURE 1.2: Contour plots for $E_1\big[\Pi_1^*(\alpha_1,\alpha_2)\big]$ (left) and $E_2\big[\Pi_2^*(\alpha_1,\alpha_2)\big]$ (right).

profits for agent i are reached when she can discriminate two categories of the same length and when the opponent can discriminate only one category (and the market size's realization is $a = 1$). When these conditions are satisfied, we have

$$\bar{\pi}_1 = \pi_1^*(a = 1, \alpha_1 = .5, \alpha_2 = 1) = 15.79$$

for agent 1, and similarly

$$\bar{\pi}_2 = \pi_2^*(a = 1, \alpha_1 = 0, \alpha_2 = .5) = 15.79$$

for agent 2.[10]

In our setting, categories are exogenous: agents are born with their threshold and they cannot change it. In [Heller and Winter, 2016], before deciding their actions, players adopt a rule rationality, i.e. a rule that determines which different situation are bundled together. Tailoring this idea to our context, we allow agents to choose how to partition the market size: in other words, agents have full control on their own thresholds. In the following paragraph, we adopt Heller and Winter's assumption and look for the equilibrium thresholds chosen when agents have full control.

As shown in Figure 1.2, agent 1 is better off in the point $\Big(\alpha_1 = 0.5, \alpha_2 = 1\Big)$, whereas agent 2 would prefer to be in the point $\Big(\alpha_1 = 0, \alpha_2 = 0.5\Big)$. Hence, each agent $i = 1, 2$ would initially choose the threshold $\alpha_i = 0.5$. However, $\Big(\alpha_1 = \alpha_2 = 0.5\Big)$ is not an equilibrium solution. In fact, when the opponent j sets $\alpha_j = 0.5$, player i prefers to deviate from $\alpha_i = 0.5$.

The problem is formalized as follows: in order to find equilibrium thresholds, we let each agent i maximize her expected profit $E_i\Big[\Pi_i^*(\alpha_1, \alpha_2)\Big]$ with respect to threshold α_i, knowing that at the same time the opponent j is making an analogous decision. In equilibrium the agents, as proved in Appendix A, set their thresholds as follows:

$$\Big(\alpha_1 = 0.415, \alpha_2 = 0.585\Big)$$

[10]The fact that realized payoffs under binary discrimination can overcome realized payoffs under perfect (and no) discrimination does not imply that binary discrimination is always beneficial with respect to perfect discrimination. In fact, when agent i realizes the maximum possible payoff $\bar{\pi}_i$, he is playing against an opponent who is not discriminating at all, because her threshold is at the boundary. Competitions between perfectly discriminating agents and between a binary categorizer and a non-discriminating agent, infact, are not comparable: the only thing that we can infer is that, *ceteris paribus*, a player may benefit from higher discrimination levels.

With respect to the initial point $\Big(\alpha_1 = \alpha_2 = 0.5\Big)$, agent 1 lowers his threshold, while agent 2 raises it. As agents set different thresholds, they ease the competition in each situation by specializing in different sectors of the market. In this case, agent 1 fine-tunes his ability to low market sizes, whereas agent 2 fine-tunes her ability to large market sizes.

This individually rational choice of thresholds is not efficient. In fact, there exists a region in the square $(0,1)^2$ where both agents have higher expected profits compared to those achieved in $\Big(\alpha_1 = 0.415, \alpha_2 = 0.585\Big)$. Agents endowed with thresholds included in this region have higher expected profits compared to agents with full control on their thresholds. To find this region, we study where the following condition is satisfied in $(0,1)^2$.

$$\begin{cases} E_1\Big[\Pi_1^*(\alpha_1,\alpha_2)\Big] \geqslant E_1\Big[\Pi_1^*(\alpha_1 = 0.415, \alpha_2 = 0.585)\Big] \\ E_2\Big[\Pi_2^*(\alpha_1,\alpha_2)\Big] \geqslant E_2\Big[\Pi_1^*(\alpha_1 = 0.415, \alpha_2 = 0.585)\Big] \end{cases} \tag{1.6}$$

We find an area, in the north-west region of $\Big(\alpha_1 = 0.415, \alpha_2 = 0.585\Big)$, under which the system (1.6) is satisfied: each of these dyads outperforms (both individually and jointly) a pair of agents that choose their thresholds.

Proposition 1.2. *There exist dyads of agents with exogenous thresholds that outperform dyads of agents that have full control on their thresholds.*

Proof. See Appendix A. □

Figure 1.3 helps visualize the question.

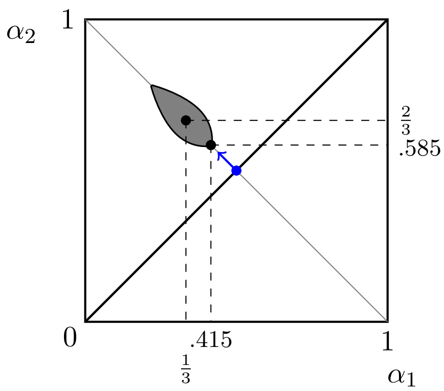


FIGURE 1.3: Agents with exogenous thresholds outperform agents with full control.

Agents that have full control on their thresholds choose $\Big(\alpha_1 = 0.415, \alpha_2 = 0.585\Big)$, as in Figure 1.3. The blue arrow represents the shift from the initial choice $(0.5, 0.5)$. All agents that are born with thresholds included in the grey area earn higher expected profits than the dyads with full control.

In fact, the endogenous choice of thresholds is very similar to a prisoners' dilemma. Each agent increases her expected profits by slightly moving her threshold: agent 1 by lowering it, and agent 2 by increasing it. On the one hand, this would increasingly mitigate competition. On the other hand, the opponent may choose a more central threshold and defect. Hence, agents will not change their thresholds, fearing the move of the opponent.

We derive two important conclusions. The first one is that, in general, the best dyads of agents have thresholds on (or near) the point $\left(\alpha_1 = \frac{1}{3}, \alpha_2 = \frac{2}{3}\right)$. In fact, when $\left(\alpha_1 = \frac{1}{3}, \alpha_2 = \frac{2}{3}\right)$, each agent is in a perfectly balanced position: on the one hand, she encounters a mild competition (milder, for instance, than agents with equal thresholds[11]); on the other hand, she has a relatively good discrimination of the market size (and, most importantly, good enough to thoroughly fit with the opponent).

In fact, the best performing agents are not the ones that discriminate R_A in two perfectly equal intervals ($a_1 = a_2 = 0.5$), but the ones with categories that best fit those of the opponents.

The second point is that expected profits only depend on thresholds (for a given a). If we put this in a long run perspective, *ex ante* expected profits found in this section suggest that there are (i) firms that consistently outperform the opponents, and (ii) dyads of firms that outperform other dyads of firms; and these differences are only driven by decision makers' discrimination abilities. This topic is more deeply analyzed in Chapter 3, where we approach the question investigating how categorization, along with the way in which a firm is organized, may affect decision makers' choices and hence firms' expected profits.

1.4 Communication

In this section, we expand the analysis allowing agents to exchange messages regarding the situation they face, and we study whether communication can increase agents' expected payoffs. As previously noted, we assume that communication is truthful, and that this is common knowledge (i.e. agents trust the message they receive). This assumption is justified by the fact that we want to analyze the overall effects of a truthful exchange of messages, rather than the actual incentives to do so.

We first analyze bilateral communication and find that there are no conditions under which it can be beneficial for both agents. Then, we analyze unilateral communication; in this case, on the contrary, such conditions exist.

1.4.1 Bilateral communication

A message $M_i = \{M_i^L, M_i^H, \emptyset\}$ is sent by agent i to let the opponent j know which situation i is facing. For instance, if $M_1 = M_1^L$, agent 1 is telling 2 that she perceives a L situation, i.e. that the market size is low.

Communication not only may refine agents' understanding of the market size, but also homogenizes their expectations. Before communication, agents individually perceive two different situations. For instance, if $a \in (\alpha_1, \alpha_2)$, agent 1 perceives the H situation, while agent 2 perceives the L situation. As soon as they exchange messages, agents mutually perceive the same combination of situations ($S_1^H \cap S_2^L$). Agent 1 updates his expectation for the market size from $E_1\left[A|S_1^H\right] = \frac{\alpha_1+1}{2}$ to $E_1\left[A|S_1^H \cap M_2^L\right] = \frac{\alpha_1+\alpha_2}{2}$. Similarly, agent 2 updates her expectation from $E_2\left[A|S_2^L\right] = \frac{\alpha_2}{2}$ to $E_2\left[A|S_2^L \cap M_1^H\right] = \frac{\alpha_1+\alpha_2}{2}$.

In this combination of situations (S_1^H and S_2^L), both agents refine their understanding of A. In the other two combinations (S_1^L and S_2^L, S_1^H and S_2^H), only one agent refines his

[11]Who are more accurately analyzed in Appendix B.

understanding (2 and 1, respectively).[12] In fact, agent $i = 1, 2$ refines her understanding of the market size only if $M_j \cap S_i^a \neq \emptyset$ and $M_j \not\supseteq S_i^a$.

As they end up with the same expectation about A, in equilibrium agents set the same production quantities. We derive the optimal production strategies for each possible combination of situations. Combining (1.2) and (1.3) and recalling that (as messages are truthful) $M_i = S_i^a$, the *interim* expected profits for agent i after perceiving the situation S_i^k and receiving the message M_j^q are

$$E_i\Big[\Pi_i|S_i^k, S_j^q\Big] = \max\Big\{x_i\Big(E_i\Big[A|S_i^k \cap S_j^q\Big] - (x_i + E_i\Big[x_j|S_i^k, S_j^q\Big])\Big), 0\Big\} \tag{1.7}$$

where $k, q \in \{L, H\}$ and $E_i\Big[x_j|S_i^k, S_j^q\Big] = x_i$. The maximization of *interim* expected payoffs in (1.7) yields three systems of two equations (3 combinations of situations $\times$ 2 agents) whose solutions are the equilibrium production functions under truthful communication. Let $x_i^{k,q*}$ denote the quantity produced by agent i when she observes S_i^k and receives the message M_j^q. Let $E_i\Big[\Pi_i^{C*}|S_1^k, S_2^q\Big]$ denote the *interim* equilibrium expected payoffs, and let $E_i\Big[\Pi_i^{C*}(\alpha_1, \alpha_2)\Big]$ denote the *ex ante* equilibrium expected payoffs.

As $E_1\Big[\Pi_1^{C*}(\alpha_1, \alpha_2)\Big] = E_2\Big[\Pi_2^{C*}(\alpha_1, \alpha_2)\Big]$ for any given pair of thresholds (α_1, α_2), differently from the case with no communication, the pair of thresholds that guarantees the highest individual *ex ante* equilibrium expected payoffs is unique. This pair is $\Big(\alpha_1 = \frac{1}{3}, \alpha_2 = \frac{2}{3}\Big)$: the dyad endowed with these thresholds expects higher payoffs compared to all other communicating dyads.

Thresholds $\Big(\alpha_1 = \frac{1}{3}, \alpha_2 = \frac{2}{3}\Big)$ divide the market size in three intervals of equal length. This reinforces the idea that specialization tends to benefit competitors: in this case, agent 1 has a fine discrimination of low market sizes, whereas agent 2 has a fine discrimination of the large market sizes. Putting together their abilities, agents are able to refine their understanding of the market size.

However, communication carries a major drawback. Under no communication diverse thresholds ease the competition between agents; under communication they actually intensify it. In fact, competition under binary categorization and communication is analogous to competition operated by agents with ternary categorization and equal thresholds. When agents have equal thresholds, they perceive exactly the same situations, and this enhances the competition within each situation. Hence, communication posits a trade-off: on the one hand, agents have an advantage in categorizing the market size, as they have the possibility to perceive three situations instead of two; on the other hand, the competition in each situation increases. In fact, as reported in footnote 10 at page 11, having a higher level of discrimination can be beneficial *per se*, but it can become detrimental when the discrimination is commonly shared with the opponent.

Consider, for instance, the expected profits that two agents with thresholds $\Big(\alpha_1 = \frac{1}{3}, \alpha_2 = \frac{2}{3}\Big)$ earn under communication or under no communication. Under communication,

$$E_1\Bigg[\Pi_1^{C*}\Bigg(\alpha_1 = \frac{1}{3}, \alpha_2 = \frac{2}{3}\Bigg)\Bigg] = E_2\Bigg[\Pi_2^{C*}\Bigg(\alpha_1 = \frac{1}{3}, \alpha_2 = \frac{2}{3}\Bigg)\Bigg] = 3.6$$

[12] The combination S_1^L and S_2^H is not possible by assumption.

Under no communication,

$$E_1\left[\Pi_1^*\left(\alpha_1=\frac{1}{3},\alpha_2=\frac{2}{3}\right)\right]=E_2\left[\Pi_2^*\left(\alpha_1=\frac{1}{3},\alpha_2=\frac{2}{3}\right)\right]=3.67$$

Although the difference is small, under no communication both agents earn slightly higher profits. In Appendix A, we prove that this result extends to each dyad (α_1, α_2) in $(0,1)^2$ because the system

$$\begin{cases} E_1\left[\Pi_1^{C*}(\alpha_1,\alpha_2)\right] \geqslant E_1\left[\Pi_1^*(\alpha_1,\alpha_2)\right] \\ E_2\left[\Pi_2^{C*}(\alpha_1,\alpha_2)\right] \geqslant E_2\left[\Pi_2^*(\alpha_1,\alpha_2)\right] \end{cases}$$

has no solution in $(0,1)^2$.[13]

Thus, communication cannot help both agents at the same time, not even with the adoption of utility transfers that redistribute payoffs.

Proposition 1.3. *Under truthful bilateral communication, agent i produces quantity* $x_i^{k,q*}$ *when she perceives situation* S_i^k *and the opponent perceives* S_j^q*, where* $k, q \in \{L, H\}$*. Ex ante expected payoffs* $E_i\left[\Pi_i^{C*}(\alpha_1,\alpha_2)\right]$ *are equal for i and j for any given dyad* (α_1, α_2)*. There is no* (α_1, α_2) *for which bilateral communication is Pareto efficient, even allowing for utility transfers.*

Proof. See Appendix A. □

The exchange of truthful messages increases the discriminating ability of both agents: before communication each agent perceives only two situations; after communication, they perceive three. For instance, agent 1 perceives $\{S_1^L, S_1^H \cap S_2^L, S_2^H\}$.

As anticipated, this resembles the case of a competition between agents with equal discrimination levels ($\#C_i = 3$ for each $i = 1, 2$). This result is not different from the one found in Section 1.2 (and, more specifically, studied in Appendix B): agents with equal discrimination levels and equal thresholds may perform poorly (with respect to agents with different categories), especially when they compete against each other.

Hence, we have (indirectly) proved that two agents with ternary categorization and that share their thresholds jointly perform worse than two agents with binary categorization and with different thresholds. Essentially, in the trade-off generated by bilateral communication between the augmented discrimination level and the augmented competition, the latter harms more then the former can benefit. This results resembles the one found in [Heller and Winter, 2016], where they find that, for certain types of games, categorical rationality may imply that information has negative value: at least one player chooses to give up additional information, as she strictly loses by obtaining it, entailing that more information is not always better.

1.4.2 Unilateral communication

We now analyze unilateral communication: we assume one agent (*sender*) sends a message to the opponent (*receiver*), but the receiver returns no message to the sender. We study whether there exist conditions under which both agents can increase their payoffs, making it a Pareto efficient solution. Again, we assume that messages are truthful.

[13] Except for the limit cases $\alpha_1 = \alpha_2$ and $(\alpha_1 = 0, \alpha_2 = 1)$.

Similarly to bilateral communication, truthful unilateral communication carries a cognitive gain for the receiver: in fact, combining the situation observed by herself and the situation observed by the sender, she can refine her partition of the market size. This is not true for the sender, who has no additional information. The updating for the receiver works exactly as in bilateral communication: she refines the perception of the market size and updates her expectations for A.

Suppose, without loss of generality, that agent 1 is the receiver and agent 2 is the sender. The sender can perceive two situations, so she sets two possible actions: let $x_{2,U}^{L*}$ and $x_{2,U}^{H*}$ denote the equilibrium productions of the sender under unilateral communication. The receiver perceives three different combinations of situations: S_1^L and M_2^L, S_1^H and M_2^L, S_1^H and M_2^H. He sets a production quantity for each of these combinations: let $x_{1,U}^{L,L*}$, $x_{1,U}^{H,L*}$, and $x_{1,U}^{H,H*}$ denote the equilibrium productions of the receiver under unilateral communication for each of these combination, respectively.

Let $E_1\Big[\Pi_1^{U*}\Big|S_1^k, M_2^q\Big]$ denote the *interim* equilibrium expected payoff for the receiver when he perceives situation S_1^k and receives message M_2^q, where $k,q \in \{L,H\}$. The convex combination of *interim* equilibrium payoff, weighted by the probability of each combination, yields the *ex ante* equilibrium expected payoff, denoted $E_1\Big[\Pi_1^{U*}(\alpha_1,\alpha_2)\Big]$. Similarly, let $E_2\Big[\Pi_2^{U*}\Big|S_2^t\Big]$ denote the *interim* equilibrium payoff for the sender, and let $E_2\Big[\Pi_2^{U*}(\alpha_1,\alpha_2)\Big]$ denote her *ex ante* equilibrium expected payoff.

As showed in the previous section, there are no conditions under which bilateral communication is adopted by any pair of players with thresholds (α_1,α_2) in $(0,1)^2$. We consider whether it can be the case for unilateral communication. Is there any region of thresholds in the square $(0,1)^2$ under which unilateral communication can simultaneously benefit both agents? As the roles of sender and receiver are not fixed, we go back to a general notation and we let $i = 1,2$ be the receiver and $j = 3-i$ be the sender; we look for a pair of thresholds in $(0,1)^2$ where two conditions are satisfied at the same time:

$$\begin{cases} E_i\Big[\Pi_i^{U*}(\alpha_1,\alpha_2)\Big] \geqslant E_i\Big[\Pi_i^{*}(\alpha_1,\alpha_2)\Big] \\ E_j\Big[\Pi_j^{U*}(\alpha_1,\alpha_2)\Big] \geqslant E_j\Big[\Pi_j^{*}(\alpha_1,\alpha_2)\Big] \end{cases} \tag{1.8}$$

The first condition guarantees that the receiver is better off under unilateral communication than under no communication; the second one that the sender is better off.[14] System (1.8) is graphically solved in Figure 1.4.

The two grey triangles are regions of thresholds for which unilateral communication benefits both agents. When agent 1 is the sender, only agents with thresholds included in the south-west grey area benefit from unilateral communication. On the contrary, when agent 2 is the sender only agents with thresholds in the north-east area benefit.

However, a key condition is required for Pareto efficiency: agents need to rely on a compensation mechanism (i.e. monetary transfers) that redistributes payoffs and ensures that the joint gains from unilateral distributed are fairly apportioned among agents. In the Appendix, we sketch a possible compensation mechanism.

Proposition 1.4. *Under truthful unilateral communication, receiver i produces quantity $x_{i,U}^{k,q*}$ as she perceives situation S_i^k and the opponent perceives S_j^q, where $k,q \in \{L,H\}$. Sender j produces quantity $x_{j,U}^{k*}$ as she perceives situation S_i^k, with $k \in \{L,H\}$. Allowing*

[14] When agents are indifferent, we assume weak preference for communication.

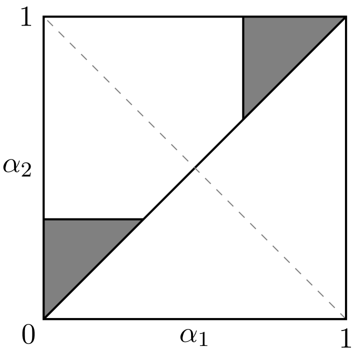


FIGURE 1.4: Region of thresholds for which unilateral communication is Pareto efficient.

for payoff transfers, there exist pairs (α_1, α_2) *for which truthful unilateral communication is Pareto efficient compared to no communication.*

Proof. See Appendix A. □

Unilateral communication is Pareto efficient only for certain types of agents; namely, those with threshold very close to the boundaries and very close to each other. These agents both discriminate very precisely the smaller (or larger) side of the market size.

Why is unilateral communication Pareto efficient, although bilateral is not? Because unilateral communication does not increase competition in each combination of situations, differently from bilateral communication. Under unilateral communication the sender only perceives two situations (exactly as under no communication) and not three as in bilateral. This eases the competition with respect to bilateral communication, yielding agents to increase their *ex ante* expected profits. In this case, communication has a positive value.

Still, the vast majority of agents would prefer not to communicate at all, especially if their thresholds α_1 and α_2 sum to one: they do not find it profitable to truthfully share their private information, hence they are not eager to start communicating.

1.5 Strategic communication

If we let agents choose the messages to send instead of assuming that they always tell the truth, they will pick them in order to maximize their profits. In this section we analyze whether agents prefer to send truthful or untruthful messages. A message is untruthful (or a *lie*) if the sender i misreports the perceived situation and tells the opponent j that she is perceiving a situation different from the one that she is actually perceiving. For instance, suppose agent i perceives the H situation. She may tell the opponent that he is actually perceiving a low situation: if j believes i, j would erroneously update her beliefs with the new (false) information and underestimate the market size. This would lead j to underproduce, leaving room for i to overproduce and earn additional profits.

We assume for simplicity that agents believe the messages that they receive, as long as they are not totally unrealistic.[15] In fact, it is possible to bundle the non-empty received messages in three types:

[15]For instance, an unrealistic message would be $M_1 = M_1^L$ when $S_2^a = S_2^H$: by construction, agent 2 knows that $S_2^a = S_2^H$ implies $S_1^a = S_1^H$.

- *Plausible.* A message is plausible if the receiver cannot tell whether it is untruthful or not. A message sent from sender i to receiver j is plausible if the following conditions are satisfied: (i) $M_i \cap S_j^a \neq \emptyset$, and (ii) $M_i \not\supseteq S_j^a$.

- *Implausible.* A message is implausible if the receiver knows with certainty that it is untruthful, or if $M_i \cap S_j^a = \emptyset$. Only untruthful messages can be implausible.

- *Predictable.* A message is predictable if the receiver knows in advance which message she is getting, or if (i) $M_i \cap S_j^a \neq \emptyset$, and (ii) $M_i \supset S_j^a$. Only truthful messages can be predictable.

Hence, we assume agents update their beliefs only with plausible or predictable messages; however, only the former ones refine their understanding of A.

There is only one rational way in which agents can lie, that is always sending the L message. In fact, there is no reason for a generic agent i to pretend that the market size is large when it is actually small, sending M_i^H when $S_i^a = S_i^L$: this lie would induce the opponent j to augment the production quantity x_j, at the point that, in order not to satiate the market, agent i would have to decrease x_i and her expected profits.

Agent 1. When agent 1 perceives the S_1^L situation, he is not willing to send untruthful messages. However, when he perceives S_1^H, he may decide to lie and issue the message $M_1 = M_1^L$. This lie is not automatically believed by agent 2. In fact, with a certain probability $\Pr(S_2^a = S_2^L) = \frac{\alpha_2 - \alpha_1}{1-\alpha_1}$ the message is plausible; with the complementary probability $\Pr(S_2^a = S_2^H) = \frac{1-\alpha_2}{1-\alpha_1}$, however, the message is implausible. If agent 2 perceives $S_2^a = S_2^H$ then she understands that the message is implausible, and will not update her beliefs. However, if agent 2 perceives $S_2^a = S_2^L$, then the message is plausible and, if she trusts agent 1, agent 2 would underproduce, leaving space for agent 1 to overproduce and earn additional profits.

Hence, agent 1 has an incentive to send an untruthful message. In the best case scenario, when $S_2^a = S_2^L$, agent 2 believes the message and this raises agent 1's expected profits. In the worst case scenario, when $S_2^a = S_2^H$, the lie is detected but nothing changes with respect to the case of truthful communication: it will be common knowledge that $a > \alpha_2$, and agents will produce $x_1^{H,H*} = x_2^{H,H*}$. As long as $\alpha_1 \leqslant \alpha_2$, there is a nonnegative probability that the lie is plausible, leading agent 1 to increase his expected profits.

Agent 2. When agent 2 decides to lie, she issues the message M_2^L when $S_2^a = S_2^H$. In this case, the message sent to agent 1 is plausible: he cannot tell if it is truthful or not. As agent 1 believes the message, he decreases the production quantity, leaving room for agent 2 to increase her quantity and, consequently, her expected profits. Thus, agent 2 has an incentive to lie as well.

As both agents prefer to send untruthful messages (i.e. send M_i^L regardless of S_i^a for each $i = 1, 2$), in equilibrium both of them, if they decide to communicate, send untruthful messages.

Proposition 1.5. *Both agents have incentives to send untruthful messages.*

Proof. See Appendix A. □

In this work we have assumed that agents only send truthful messages. However, in this section we have shown that agents prefer to send untruthful ones. As both agents know that the opponent has incentives to lie, none of them would ever be fully credible. So, unless there exists a strategyproof mechanism that forces agents to tell the truth, agents' messages carry no import. Then, the case of communication and no communication coincide, in terms of final results.

A strategyproof mechanism could be implemented if messages are verifiable. For instance, agents could sign a contract that assigns a large fine for whoever sends untruthful messages. Nonetheless, the study of strategyproof mechanisms goes beyond the scope of this paper.

1.6 Conclusions

This chapter studies the impact of categorical thinking on decision making processes in a competitive setting. Our starting point is the fact, widely recognized both in cognitive science and in the behavioral economics literature, that agents, due to their cognitive limitations, cannot fully perceive the information they get from their environment. This may induce them to pool similar information together, because memories are easier to access and the cost to store new information is reduced. However, there is a loss in the precision of the information acquired: in this sense, we say that agents are categorically rational; they are rational conditional on the boundaries of their mental categories.

We use a Cournot game to study how categorical thinking affects equilibrium strategies and expected payoffs, assuming that contiguous categories are exogenous and fixed. Then, we introduce a preliminary stage in the game where agents choose their thresholds, as in [Heller and Winter, 2016]. We find that there exists dyads of agents with exogenous thresholds that perform better (both individually and jointly) than those who can choose their categories. In fact, the choice of equilibrium thresholds resembles a prisoners' dilemma: if they both move their thresholds, agents reap higher profits because they mitigate competition. However, they also give the opponent the opportunity to defect, choosing a more central threshold. Because agents with certain thresholds can perform better than other dyads (on average), this suggests that unconscious coordination may emerge even in a competitive setting.

We analyze conditions under which communication can increase expected profits. Because of the competitive nature of the game, communication posits a trade-off between a cognitive advantage (agents improve their perception of the market size) and a competition disadvantage (an increase in the strength of the competition within each situation). We find that, under specific conditions (agents with thresholds very close to the boundaries and very close to each other, utility transfers, truthful communication), only unilateral communication can be beneficial for both agents. This result suggests that the trade-off is positively solved only when communication is partial, in the sense that only one agent sends messages. Hence, the value of communication is positive only when communication is unilateral, although it becomes negative when communication is bilateral. Loosely speaking, little information (only one agent sends messages) is good, but too much information (both exchange messages) is bad.

The limiting assumptions that we imposed imply that there exists multiple ways in which this model could be extended. The first one is in the choice of the game: our results mainly depend on the nature of the competition, which could be extended including different games (for instance, a Bertrand competition) and different assumptions ($n > 2$ players, positive marginal costs, and so on). Secondly, a more complete analysis of the message exchange game could be object of future extensions: in particular, study what would happen if we let agents be fully rational. Moreover, our results crucially hinge on the simplifying assumption that the state space is fully ordered. Other improvements are possible: for instance, the choice of the distribution of the random variable A, the number of initial categories, and agents' learning capacity.

Appendix A

Proposition 1.1. *Agent i produces the equilibrium quantity x_i^{k*} when she perceives situation S_i^k, where $i = 1, 2$ and $k \in \{L, H\}$. Ex ante expected payoffs are $E_i\left[\Pi_i^*(\alpha_1, \alpha_2)\right]$ and depend on thresholds α_1 and α_2.*

Proof. Given (1.1), we know that agent 1 knows that agent 2 perceives the situation $S_2^A = \{S_2^L, S_2^H\}$ with the following probabilities

$$\Pr\left(S_2^A = S_2^L | S_1^A = S_1^L\right) = 1,\ \Pr\left(S_2^A = S_2^H | S_1^A = S_1^L\right) = 0$$

and

$$\Pr\left(S_2^A = S_2^L | S_1^A = S_1^H\right) = \frac{\alpha_2 - \alpha_1}{1 - \alpha_1},\ \Pr\left(S_2^A = S_2^H | S_1^A = S_1^H\right) = \frac{1 - \alpha_2}{1 - \alpha_1}$$

On the other hand, agent 2 knows that agent 1 perceives the situation $S_1^A = \{S_1^L, S_1^H\}$ with the following probabilities

$$\Pr\left(S_1^A = S_1^L | S_2^A = S_2^L\right) = \frac{\alpha_1}{\alpha_2},\ \Pr\left(S_1^A = S_1^H | S_2^A = S_2^L\right) = \frac{\alpha_2 - \alpha_1}{\alpha_2}$$

and

$$\Pr\left(S_1^A = S_1^L | S_2^A = S_2^H\right) = 0,\ \Pr\left(S_1^A = S_1^H | S_2^A = S_2^H\right) = 1$$

These probabilities help us to write down the following system of best response functions

$$\begin{cases} \frac{\partial E_1\left[\Pi_1(X) | S_1^A = S_1^L\right]}{\partial x_1^L} = 0, \partial \frac{{x_1}^L\left(\alpha_1/2 - {x_1}^L - {x_2}^L\right)}{\partial x_1^L} = 0 \\ \frac{\partial E_1\left[\Pi_1(X) | S_1^A = S_1^H\right]}{\partial x_1^H} = 0, \partial \frac{{x_1}^H\left(\alpha_1/2 + 1/2 - {x_1}^H - \frac{(\alpha_2 - \alpha_1){x_2}^L}{1 - \alpha_1} - \frac{(1 - \alpha_2){x_2}^H}{1 - \alpha_1}\right)}{\partial x_1^H} = 0 \\ \frac{\partial E_2\left[\Pi_2(X) | S_2^A = S_2^L\right]}{\partial x_2^L} = 0, \partial \frac{{x_2}^L\left(\alpha_2/2 - {x_2}^L - \frac{\alpha_1\, {x_1}^L}{\alpha_2} - \frac{(\alpha_2 - \alpha_1){x_1}^H}{\alpha_2}\right)}{\partial x_2^L} = 0 \\ \frac{\partial E_2\left[\Pi_2(X) | S_2^A = S_2^H\right]}{\partial x_2^H} = 0, \partial \frac{{x_2}^H\left(\alpha_2/2 + 1/2 - {x_1}^H - {x_2}^H\right)}{\partial x_2^H} = 0 \end{cases}$$

Solving the system yields to the Nash-equilibrium production function, that can be written as:

$$\begin{aligned} x_1^{L*} &= \frac{1}{6} \frac{6\alpha_1^2\alpha_2 - 3\alpha_1\alpha_2^2 - 6\alpha_1\alpha_2 + 3\alpha_2^2 + \alpha_1 - \alpha_2}{3\alpha_1\alpha_2 + \alpha_1 - 4\alpha_2} \\ x_1^{H*} &= \frac{1}{6} \frac{6\alpha_1^2\alpha_2 - 3\alpha_1\alpha_2^2 + \alpha_1 - 4\alpha_2}{3\alpha_1\alpha_2 + \alpha_1 - 4\alpha_2} \\ x_2^{L*} &= \frac{1}{6} \frac{-3\alpha_1^2\alpha_2 + 6\alpha_1\alpha_2^2 + 3\alpha_1^2 - 6\alpha_2^2 - 2\alpha_1 + 2\alpha_2}{3\alpha_1\alpha_2 + \alpha_1 - 4\alpha_2} \\ x_2^{H*} &= \frac{1}{6} \frac{-3\alpha_1^2\alpha_2 + 6\alpha_1\alpha_2^2 + 6\alpha_1\alpha_2 - 6\alpha_2^2 + \alpha_1 - 4\alpha_2}{3\alpha_1\alpha_2 + \alpha_1 - 4\alpha_2} \end{aligned}$$

We substitute these quantities back in the respective profit functions in order to find the equilibrium *interim* payoffs, which can be written as

$$E_1\left[\Pi_1^*|S_1^L\right] = \frac{1}{36}\left(\frac{6\alpha_1^2\alpha_2 - 3\alpha_1\alpha_2^2 - 6\alpha_1\alpha_2 + 3\alpha_2^2 + \alpha_1 - \alpha_2}{3\alpha_1\alpha_2 + \alpha_1 - 4\alpha_2}\right)^2$$
$$E_1\left[\Pi_1^*|S_1^H\right] = \frac{1}{36}\left(\frac{6\alpha_1^2\alpha_2 - 3\alpha_1\alpha_2^2 + \alpha_1 - 4\alpha_2}{3\alpha_1\alpha_2 + \alpha_1 - 4\alpha_2}\right)^2$$
$$E_2\left[\Pi_2^*|S_2^L\right] = \frac{1}{36}\left(\frac{-3\alpha_1^2\alpha_2 + 6\alpha_1\alpha_2^2 + 3\alpha_1^2 - 6\alpha_2^2 - 2\alpha_1 + 2\alpha_2}{3\alpha_1\alpha_2 + \alpha_1 - 4\alpha_2}\right)^2$$
$$E_2\left[\Pi_2^*|S_2^H\right] = \frac{1}{36}\left(\frac{-3\alpha_1^2\alpha_2 + 6\alpha_1\alpha_2^2 + 6\alpha_1\alpha_2 - 6\alpha_2^2 + \alpha_1 - 4\alpha_2}{3\alpha_1\alpha_2 + \alpha_1 - 4\alpha_2}\right)^2$$

Weighting the *interim* expected payoff function by the probability that i perceives each situation ($\Pr(S_i^A = S_i^L) = \alpha_i$ and $\Pr(S_i^A = S_i^H) = 1 - \alpha_i$), we can write down the *ex ante* payoffs as

$$E_1\left[\Pi_1^*(\alpha_1,\alpha_2)\right] = \frac{-36\alpha_1^4\alpha_2^2 + 36\alpha_1^3\alpha_2^3 - 9\alpha_1^2\alpha_2^4 + 72\alpha_1^3\alpha_2^2 - 54\alpha_1^2\alpha_2^3 + 9\alpha_1\alpha_2^4 - 36\alpha_1^2\alpha_2^2}{36\left(3\alpha_1\alpha_2 + \alpha_1 - 4\alpha_2\right)^2} + $$
$$+ \frac{18\alpha_1\alpha_2^3 + 6\alpha_1^2\alpha_2 - 15\alpha_1\alpha_2^2 + \alpha_1^2 - 8\alpha_1\alpha_2 + 16\alpha_2^2}{36\left(3\alpha_1\alpha_2 + \alpha_1 - 4\alpha_2\right)^2}$$
$$E_2\left[\Pi_2^*(\alpha_1,\alpha_2)\right] = \frac{-9\alpha_1^4\alpha_2^2 + 36\alpha_1^3\alpha_2^3 - 36\alpha_1^2\alpha_2^4 + 9\alpha_1^4\alpha_2 - 18\alpha_1^3\alpha_2^2 - 36\alpha_1^2\alpha_2^3 + 72\alpha_1\alpha_2^4 - 18\alpha_1^3\alpha_2}{36\left(3\alpha_1\alpha_2 + \alpha_1 - 4\alpha_2\right)^2} + $$
$$+ \frac{72\alpha_1^2\alpha_2^2 - 36\alpha_1\alpha_2^3 - 36\alpha_2^4 + 15\alpha_1^2\alpha_2 - 60\alpha_1\alpha_2^2 + 36\alpha_2^3 + \alpha_1^2 - 8\alpha_1\alpha_2 + 16\alpha_2^2}{36\left(3\alpha_1\alpha_2 + \alpha_1 - 4\alpha_2\right)^2}$$

Once the market size is realized, agents earn the following realized profits:

$$\pi_1^*(a,\alpha_1,\alpha_2) = \begin{cases} \max\left\{x_1^{L*}\left(a - (x_1^{L*} + x_2^{L*})\right), 0\right\} & \text{if } a \in (0,\alpha_1) \\ \max\left\{x_1^{H*}\left(a - (x_1^{H*} + \frac{\alpha_2-\alpha_1}{1-\alpha_1}x_2^{L*} + \frac{1-\alpha_2}{1-\alpha_1}x_2^{H*})\right), 0\right\} & \text{if } a \in (\alpha_1,1) \end{cases}$$

for agent 1, and

$$\pi_2^*(a,\alpha_1,\alpha_2) = \begin{cases} \max\left\{x_2^{L*}\left(a - (\frac{\alpha_1}{\alpha_2}x_1^{L*} + \frac{\alpha_2-\alpha_1}{\alpha_2}x_1^{H*} + x_2^{L*})\right), 0\right\} & \text{if } a \in (0,\alpha_2) \\ \max\left\{x_2^{H*}\left(a - (x_1^{H*} + x_2^{H*})\right), 0\right\} & \text{if } a \in (\alpha_2,1) \end{cases}$$

for agent 2. □

Proposition 1.2. *There exist dyads of agents with exogenous thresholds that outperform dyads of agents that have full control on their thresholds.*

Proof. We first prove that $(\alpha_1,\alpha_2) = (0.415, 0.585)$ effectively represents the pair of thresholds chosen by agents 1 and 2 if they had control on their thresholds. In order to find the equilibrium pair of thresholds, we let each agent i maximize his/her expected profits $E_i\left[\Pi_i^*(\alpha_1,\alpha_2)\right]$ with respect to his threshold α_i:

$$\begin{cases} \max_{\alpha_1} \; E_1\left[\Pi_1^*(\alpha_1,\alpha_2)\right] \\ \max_{\alpha_2} \; E_2\left[\Pi_2^*(\alpha_1,\alpha_2)\right] \end{cases}$$

Assuming $\alpha_2 \neq 0$, $\alpha_1 \neq 1$, $\alpha_2 \neq 2\alpha_1 - 1$, $\alpha_2 \neq \frac{\alpha_1}{2}$, and $\alpha_1 \neq \frac{3\alpha_2+\alpha_2}{3\alpha_2+1}$, and then collecting α_2 in the first equation and α_1 in the second, it is possible to rewrite the system as:

$$\begin{cases} (5\alpha_1 - 4)\alpha_2^2 + (12\alpha_1^3 - 38\alpha_1^2 + 28\alpha_1 - 4)\alpha_2 + (4\alpha_1^3 - 2\alpha_1^2 - \alpha_1) = 0 \\ (5\alpha_2 - 1)\alpha_1^2 + (-12\alpha_2^3 - 2\alpha_2^2 + 2\alpha_2)\alpha_1 + (16\alpha_2^3 - 8\alpha_2^2) = 0 \end{cases}$$

Call the first equation $eq_1(\alpha_1, \alpha_2) = 0$ and the second equation $eq_2(\alpha_1, \alpha_2) = 0$. Rewrite $eq_2(\alpha_1, \alpha_2) = 0$ as $a \cdot \alpha_1^2 + b \cdot \alpha_1 + c = 0$. In order to solve it as a second degree equation, a, b and c must be different from zero, hence $\alpha_2 \neq \{\frac{1}{5}, 0, -\frac{1}{2}, \frac{1}{3}, \frac{1}{2}\}$. The discriminant of the equation is $\Delta = 4\alpha_2^2(36\alpha_2^4 + 12\alpha_2^3 - 91\alpha_2^2 + 54\alpha_2 - 7)$, and it is strictly positive in the following intervals: $\alpha_2 \in (0.185, 0.591)$ and $\alpha_2 \in (0.889, 1]$.

Hence, we can write the two roots $\alpha_{1,1}$ and $\alpha_{1,2}$:

$$\alpha_{1,1} = \frac{-b + \sqrt{\Delta}}{4a} = \frac{\alpha_2(6\alpha_2^2 + \alpha_2 - 1 + \sqrt{36\alpha_2^4 + 12\alpha_2^3 - 91\alpha_2^2 + 54\alpha_2 - 7})}{5\alpha_2 - 1}$$
$$\alpha_{1,2} = \frac{-b - \sqrt{\Delta}}{4a} = \frac{\alpha_2(6\alpha_2^2 + \alpha_2 - 1 - \sqrt{36\alpha_2^4 + 12\alpha_2^3 - 91\alpha_2^2 + 54\alpha_2 - 7})}{5\alpha_2 - 1}$$

Now, we substitute these solutions back in the first equation and numerically solve it, as it only depends on α_2. Hence, we look for the roots of $eq_1(\alpha_{1,1}, \alpha_2) = 0$ and $eq_2(\alpha_{1,2}, \alpha_2) = 0$.

$$eq_1(\alpha_{1,1}, \alpha_2) = 0 \Rightarrow \alpha_{2,1} = 0.209 \vee \alpha_{2,2} = 0.333$$
$$eq_1(\alpha_{1,2}, \alpha_2) = 0 \Rightarrow \alpha_{2,3} = 0 \vee \alpha_{2,4} = 0.585$$

Solution $\alpha_{2,3} = 0$ is not admissible. Let us check the other ones substituting $\alpha_{2,1}$ and $\alpha_{2,2}$ back in $\alpha_{1,1}$, and $\alpha_{2,4}$ back in $\alpha_{1,2}$.

$$\alpha_{1,1}(\alpha_{2,1}) = 0.729$$
$$\alpha_{1,1}(\alpha_{2,2}) = 0.667$$
$$\alpha_{1,2}(\alpha_{2,4}) = 0.415$$

We have three candidate solutions: $(\alpha_1 = 0.729, \alpha_2 = 0.209)$, $(\alpha_1 = 0.667, \alpha_2 = 0.333)$ and $(\alpha_1 = 0.415, \alpha_2 = 0.585)$. The first two candidate solutions are not admissible (we assumed $\alpha_1 \leqslant \alpha_2$). Consequently, the only candidate solution which is admissible is $(\alpha_1 = 0.415, \alpha_2 = 0.585)$.

Alternatively, the system can be graphically solved as in the following Figure 1.5. The steepest curve is obtained by plotting the solution of the first equation for α_1 when α_2 is a sequence of values between 0 and 1 taken every .005 steps (i.e, $\alpha_2 = 0, 0.005, 0.01, ..., 1$). Being a quartic equation, there are four roots: two of them are imaginary, and another one falls out of the feasible region for α_1. The same process has been applied to the second equation, fixing α_1 and solving for α_2: plotting the solution we are able to sketch the other curve.

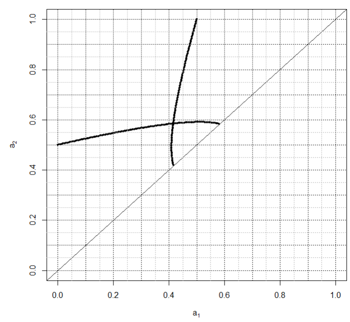


FIGURE 1.5: Graphical resolution of equilibrium thresholds

We now show that there exists a locus of points in $\alpha_1 \times \alpha_2$ under which the *ex ante* expected payoff is higher than the one computed in $(\alpha_1 = 0.415, \alpha_2 = 0.585)$. In other words, we graphically show the solution of system (1.6) in the Figure 1.6 (we added the assumption that $\alpha_1 \leqslant \alpha_2$).

The grey area in the north-western triangle of the square represents the pairs of thresholds that satisfy condition (1.6). Note that what happens in the other triangle (where $\alpha_1 > \alpha_2$) does not belong to the feasible area. □

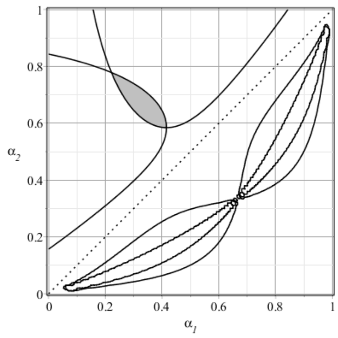


FIGURE 1.6: Area in which agents with exogenous thresholds outperform agents with full control.

Proposition 1.3. *Under truthful bilateral communication, agent i produces quantity $x_i^{k,q*}$ when she perceives situation S_i^k and the opponent perceives S_j^q, where $k, q \in \{L, H\}$. Ex ante expected payoffs $E_i\left[\Pi_i^{C*}(\alpha_1, \alpha_2)\right]$ are equal for i and j for any given dyad (α_1, α_2). There is no (α_1, α_2) for which bilateral communication is Pareto efficient, even allowing for utility transfers.*

Proof. Maximizing *ex post* expected payoffs $E_i\left[\Pi_i(X)\middle|S_i^t, S_i^s\right]$ in (1.7) we find the following equilibrium production function for $i = 1, 2$ and $t, s \in \{L, H\}$:

$$x_i^{t,s*} = \frac{E_i\left[A|S_i^t \cap S_j^s\right]}{3}$$

where $E_i\left[A|S_i^L \cap S_j^L\right] = \frac{\alpha_1}{2}$, $E_i\left[A|S_i^H \cap S_j^L\right] = \frac{\alpha_1+\alpha_2}{2}$, and $E_i\left[A|S_i^H \cap S_j^H\right] = \frac{\alpha_2+1}{2}$. *Ex post* expected payoffs are

$$E_i\left[\Pi_i^{C*}|S_1^L \cap S_2^L\right] = \frac{\alpha_1^2}{36}$$
$$E_i\left[\Pi_i^{C*}|S_1^H \cap S_2^L\right] = \frac{(\alpha_1+\alpha_2)^2}{36}$$
$$E_i\left[\Pi_i^{C*}|S_1^H \cap S_2^H\right] = \frac{(\alpha_2+1)^2}{36}$$

and *ex ante* expected payoffs are

$$E_i\left[\Pi_i^{C*}(\alpha_1, \alpha_2)\right] = \frac{-\alpha_1^2\alpha_2 + \alpha_1\alpha_2^2 - \alpha_2^2 + \alpha_2 + 1}{36}$$

for both $i = 1, 2$. To find the pair (α_1, α_2) that maximizes $E_i\left[\Pi_i^{C*}(\alpha_1, \alpha_2)\right]$, we solve the following maximization problem:

$$\max_{\alpha_1, \alpha_2} \; E_i\left[\Pi_i^{C*}(\alpha_1, \alpha_2)\right]$$

The solution is exactly the pair $\left(\alpha_1 = \frac{1}{3}, \alpha_2 = \frac{2}{3}\right)$.

We now prove that communication cannot benefit both agents at the same time. To do so, we compare *ex ante* expected payoffs from truthful communication with *ex ante* expected payoffs from no communication. We say that agent i benefits from communication if

$$E_i\left[\Pi_i^{C*}(\alpha_1, \alpha_2)\right] \geqslant E_i\left[\Pi_i^*(\alpha_1, \alpha_2)\right]$$

These two condition (one for each agent) are plotted in the following Figure 1.7.

In the left (right) square, the grey area represents the locus of points where the profitability condition is satisfied for agent 1 (agent 2). In general, agents benefit from communication when their threshold is relatively less central than the opponent's one. Moreover, if we superimpose the two squares we find that

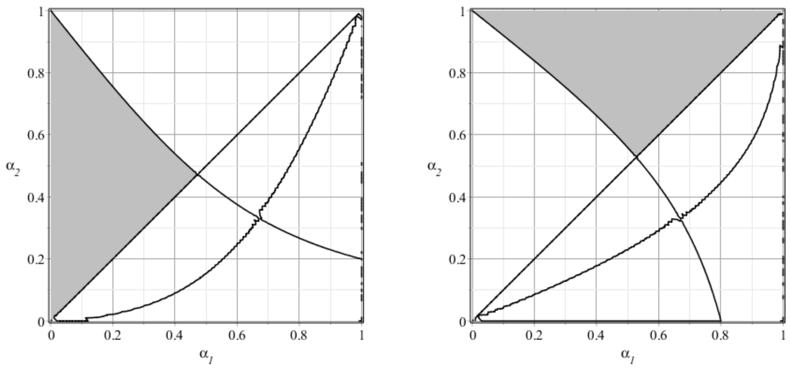


FIGURE 1.7: Profitability of truthful communication: agent 1 (left) and agent 2 (right).

there is not grey area in common, meaning that there are no dyads of agents that have incentives to use bilateral communication. In other words, we have that the system

$$\begin{cases} E_1\left[\Pi_1^{C*}(\alpha_1,\alpha_2)\right] \geqslant E_1\left[\Pi_1^{*}(\alpha_1,\alpha_2)\right] \\ E_2\left[\Pi_2^{C*}(\alpha_1,\alpha_2)\right] \geqslant E_2\left[\Pi_2^{*}(\alpha_1,\alpha_2)\right] \end{cases}$$

has no solution for α_1 and α_2 included in $(0,1)^2$, except for two cases in which these profits are equal, meaning that communication is useless: (i) $\alpha_1 = \alpha_2$, and $\big(\alpha_1 = 0, \alpha_2 = 1\big)$. We now prove that, even allowing for utility transfers, bilateral communication would still not be Pareto efficient. In fact, the sum of agents' expected profits under bilateral communication is, for all $(\alpha_1,\alpha_2) \in (0,1)^2$, smaller than the sum of agents' expected profits under no communication. In other words, the condition

$$E_1\left[\Pi_1^{C*}(\alpha_1,\alpha_2)\right] + E_2\left[\Pi_2^{C*}(\alpha_1,\alpha_2)\right] \geqslant E_1\left[\Pi_1^{*}(\alpha_1,\alpha_2)\right] + E_2\left[\Pi_2^{*}(\alpha_1,\alpha_2)\right]$$

has no solution in the region $(\alpha_1,\alpha_2) \in (0,1)^2$. As the sum of expected payoffs under communication is smaller than the sum under no communication, it is not possible that, even with utility transfers, communication makes one agent better off and the other one indifferent. Hence, bilateral communication is not Pareto efficient. □

Proposition 1.4. *Under truthful unilateral communication, receiver i produces quantity $x_{i,U}^{k,q*}$ as she perceives situation S_i^k and the opponent perceives S_j^q, where $k,q \in \{L,H\}$. Sender j produces quantity $x_{j,U}^{k*}$ as she perceives situation S_i^k, with $k \in \{L,H\}$. Allowing for payoff transfers, there exist pairs (α_1,α_2) for which truthful unilateral communication is Pareto efficient compared to no communication.*

Proof. To sketch the proof, we first find the expected profits of agents 1 and 2 with unilateral communication (agent 2 is the sender of the truthful message, agent 1 is the receiver).

After perceiving the situation $S_2^A = S_2^H$, agent 2 solves the following maximization problem

$$\max_{x_{2,U}^H} \; E_2[\Pi_2^{U*}|S_2^H] = \big(E_2[A|S_2^H] - (x_{1,U}^{H,H} + x_{2,U}^H)\big)\cdot x_2^H$$

where $x_{1,U}^{H,L*}$ the quantity produced by agent 1 when $S_1^A = S_1^H$ and $S_2^A = M_2^A = S_2^L$. Agent 2 knows for sure that, when $S_2^A = S_2^H$, also $S_1^A = S_1^H$. The best reply follows from the first order condition of the problem:

$$x_{2,U}^H = \frac{E_2[A|S_2^H] - x_{1,U}^{H,H}}{2}$$

After perceiving the situation $S_2^A = S_2^L$, agent 2 knows there is a positive probability that agent 1 perceives $S_1^A = S_1^L$, and a complementary probability that he perceives $S_1^A = S_1^H$. Hence, the maximization problem becomes:

$$\max_{x^L_{2,U}} E_2[\Pi^U_2|S^L_2] =$$
$$= \Pr(S^A_1 = S^L_1|S^A_2 = S^L_2) \cdot E_2[\Pi^U_2|S^L_2, S^L_1] + \Pr(S^A_1 = S^H_1|S^A_2 = S^L_2) \cdot E_2[\Pi^U_2|S^L_2, S^H_1]$$
$$= \left(E_2[A|S^L_2] - \left(x^L_{2,U} + \frac{\alpha_1}{\alpha_2}x^{L,L}_{1,U} + \frac{\alpha_2-\alpha_1}{\alpha_2}x^{H,L}_{1,U}\right)\right) \cdot x^L_{2,U}$$

The best response function becomes:

$$x^L_{2,U} = \frac{1}{2}\left(E_2[A|S^L_2] - \left(\frac{\alpha_1}{\alpha_2}x^{L,L}_{1,U} + \frac{\alpha_2-\alpha_1}{\alpha_2}x^{L,H}_{1,U}\right)\right)$$

When agent 1 perceives $S^A_1 = S^L_1$, and at the same time observes that $M^A_2 = M^L_2$, he realizes that agent 2 plays $x^L_{2,U}$, and then his best reply function will be

$$x^{L,L}_{1,U} = \frac{E_1[A|S^L_1 \cap S^L_2] - x^L_{2,U}}{2}$$

When $S^A_1 = S^H_1$, agent 1 plays $x^{H,L}_{1,U}$ if $M^A_2 = M^L_2$ and $x^{H,H}_{1,U}$ if $M^A_2 = M^H_2$:

$$x^{H,L}_{1,U} = \frac{E_1[A|S^H_1 \cap S^L_2] - x^L_{2,U}}{2}$$
$$x^{H,H}_{1,U} = \frac{E_1[A|S^H_1 \cap S^H_2] - x^H_{2,U}}{2}$$

Thus, in order to find the equilibrium strategies we need to solve a system of five equations in five unknowns:

$$\begin{cases} x^L_{2,U} = \frac{1}{2}\left(E_2[A|S^L_2] - \left(\frac{\alpha_1}{\alpha_2}x^{L,L}_{1,U} + \frac{\alpha_2-\alpha_1}{\alpha_2}x^{L,H}_{1,U}\right)\right) \\ x^H_{2,U} = \frac{E_2[A|S^H_2]-x^{H,H}_{1,U}}{2} \\ x^{L,L}_{1,U} = \frac{E_1[A|S^L_1 \cap S^L_2]-x^L_{2,U}}{2} \\ x^{H,L}_{1,U} = \frac{E_1[A|S^H_1 \cap S^L_2]-x^L_{2,U}}{2} \\ x^{H,H}_{1,U} = \frac{E_1[A|S^H_1 \cap S^H_2]-x^H_{2,U}}{2} \end{cases}$$

where $E_2[A|S^L_2] = \frac{\alpha_2}{2}$, $E_2[A|S^H_2] = \frac{1+\alpha_2}{2}$, $E_1[A|S^L_1 \cap S^L_2] = \frac{\alpha_1}{2}$, $E_1[A|S^H_1 \cap S^L_2] = \frac{\alpha_1+\alpha_2}{2}$, and $E_1[A|S^H_1 \cap S^H_2] = \frac{1+\alpha_2}{2}$. Equilibrium quantities are described hereafter:

$$\begin{cases} x^{L*}_{2,U} = \frac{\alpha_2}{6} \\ x^{H*}_{2,U} = \frac{1+\alpha_2}{6} \\ x^{L,L*}_{1,U} = \frac{3\alpha_1-\alpha_2}{12} \\ x^{H,L*}_{1,U} = \frac{2\alpha_1+3\alpha_2}{12} \\ x^{H,H*}_{1,U} = \frac{1+\alpha_2}{6} \end{cases}$$

Substituting these optimal quantities back in the respective profit functions, we find the equilibrium profits conditional on the situations perceived by each agent.

$$E_1\left[\Pi_1^{U*}|S_1^L \cap S_2^L\right] = \frac{(-\alpha_2 + 3\,\alpha_1)^2}{144}$$
$$E_1\left[\Pi_1^{U*}|S_1^L \cap S_2^H\right] = \frac{(2\,\alpha_2 + 3\,\alpha_1)^2}{144}$$
$$E_1\left[\Pi_1^{U*}|S_1^H \cap S_2^H\right] = \frac{(-\alpha_2 + 3\,\alpha_1)^2}{144}$$
$$E_2\left[\Pi_2^{U*}|S_2^L\right] = \frac{\alpha_2^2}{36}$$
$$E_2\left[\Pi_2^{U*}|S_2^H\right] = \frac{(\alpha_1 + 1)^2}{36}$$

Weighting these *interim* expected profits by the probability of each situation (or combination of situations) we find

$$E_1\left[\Pi_1^{U*}(\alpha_1,\alpha_2)\right] = \frac{-\alpha_1^2\alpha_2}{16} + \frac{\alpha_1\alpha_2^2}{16} - \frac{\alpha_2^2}{36} + \frac{\alpha_2}{36} + \frac{1}{36}$$

and

$$E_2\left[\Pi_2^{U*}(\alpha_1,\alpha_2)\right] = \frac{1}{36}\Big(-\alpha_2^2 + \alpha_2 + 1\Big)$$

We now prove that unilateral communication cannot benefit both agents at the same time. To do so, we compare *ex ante* expected payoffs from unilateral truthful communication with *ex ante* expected payoffs from no communication. We say that agent i benefits from unilateral communication if

$$E_i\left[\Pi_i^{U*}(\alpha_1,\alpha_2)\right] \geqslant E_i\left[\Pi_i^{*}(\alpha_1,\alpha_2)\right]$$

These two condition (one for each agent) are plotted in Figure 1.8.

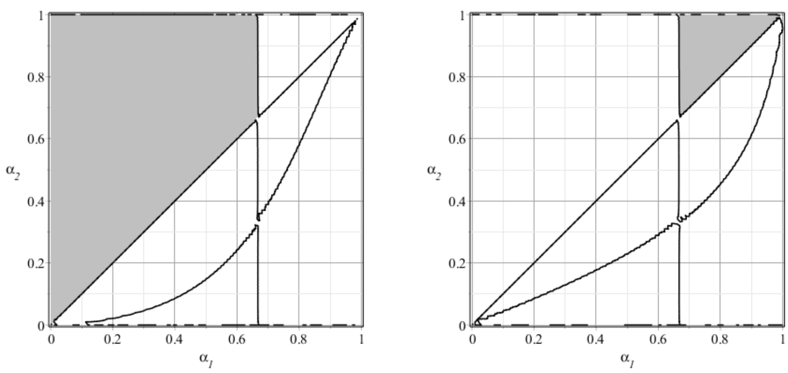


FIGURE 1.8: Profitability of truthful unilateral communication: agent 1 (left) and agent 2 (right).

We now prove that there exists an area in the square $(\alpha_1,\alpha_2) \in (0,1)^2$ under which the sum of agents' expected payoffs under unilateral communication is greater than the sum of expected payoffs under no communication. The condition

$$E_1\left[\Pi_1^{U*}(\alpha_1,\alpha_2)\right] + E_2\left[\Pi_2^{U*}(\alpha_1,\alpha_2)\right] \geqslant E_1\left[\Pi_1^{*}(\alpha_1,\alpha_2)\right] + E_2\left[\Pi_2^{*}(\alpha_1,\alpha_2)\right]$$

is plotted in Figure 1.9.

In the grey area, utility transfers help achieving a Pareto efficient solution.

Define $c(l_{ij}, t_i^*, t_j^*)$ a two-step *compensation mechanism* (i.e. a utility transfer) that depends on two variables: (i) the amount of profits l_{ij} that i needs to transfer to j in order to make the latter at least indifferent between the new setting and the old one; (ii) (t_i^*, t_j^*), the apportionment under Nash bargaining of what is left after the first compensation l_{ij}.

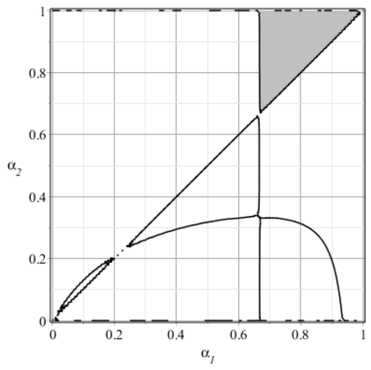


FIGURE 1.9: Area of (α_1, α_2) under which unilateral communication can be Pareto efficient.

At an individual level, agent i has a *gain* (*loss*) from unilateral communication if the difference $E_i\left[\Pi_i^{\mathrm{u}}(X)\right] \geqslant E_i\left[\Pi_i(X)\right]$ is positive (negative). At the aggregate level, there exists a *social gain* (*social loss*) from unilateral communication if the difference

$$\left(E_1\left[\Pi_1^{U*}(\alpha_1,\alpha_2)\right] + E_2\left[\Pi_2^{U*}(\alpha_1,\alpha_2)\right]\right) - \left(E_1\left[\Pi_1^{*}(\alpha_1,\alpha_2)\right] + E_2\left[\Pi_2^{*}(\alpha_1,\alpha_2)\right]\right)$$

is positive (negative).

We seek conditions under which a compensation mechanism can be applied to reach a Pareto improvement. A Pareto improvement is formally defined by the following system:

$$\begin{cases} E_i\left[\Pi_i^{U*}(\alpha_1,\alpha_2)\right] > E_i\left[\Pi_i^{*}(\alpha_1,\alpha_2)\right] \\ E_j\left[\Pi_j^{U*}(\alpha_1,\alpha_2)\right] \geqslant E_j\left[\Pi_j^{*}(\alpha_1,\alpha_2)\right] \end{cases} \tag{1.9}$$

For a Pareto improvement to exist, both agents need to be at least better off and one of them (in (1.9), agent i) needs to be strictly better off. Conditions in system (1.9) imply the following one:

$$E_i\left[\Pi_i^{U*}(\alpha_1,\alpha_2)\right] + E_j\left[\Pi_j^{U*}(\alpha_1,\alpha_2)\right] > E_i\left[\Pi_i^{*}(\alpha_1,\alpha_2)\right] + E_j\left[\Pi_j^{*}(\alpha_1,\alpha_2)\right] \tag{1.10}$$

which is exactly the definition of *social gain*; hence, a *social gain* is necessary in order to have a Pareto improvement (whereas it is possible to have a *social gain* without having a Pareto improvement). This is a necessary condition for the compensation mechanism $c(l_{ij}, t_i, t_j)$ to be applicable: in fact, the first transfer l_{ij} from i to j is possible only if the relative *gain* of agent i is greater then the relative *loss* of agent j. In fact, condition (1.10) can be rewritten as:

$$E_i\left[\Pi_i^{U*}(\alpha_1,\alpha_2)\right] - E_i\left[\Pi_i^{*}(\alpha_1,\alpha_2)\right] > E_j\left[\Pi_j^{*}(\alpha_1,\alpha_2)\right] - E_j\left[\Pi_j^{U*}(\alpha_1,\alpha_2)\right]$$

If there is a *social gain*, the amount l_i is the money transfer from agent i to agent j that makes the latter indifferent. Hence, we have

$$l_i = E_j\left[\Pi_j^{*}(\alpha_1,\alpha_2)\right] - E_j\left[\Pi_j^{U*}(\alpha_1,\alpha_2)\right]$$

Let us now analyze the second step. Call T_i the net gain of agent i after he has compensated agent j for the loss of utility. We have that

$$T_i = E_i\left[\Pi_i^{U*}(\alpha_1,\alpha_2)\right] - E_i\left[\Pi_i^{*}(\alpha_1,\alpha_2)\right] - l_i$$

Condition (1.10) guarantees that T_i is strictly positive. Let t_i^* and t_j^* be the result of symmetric Nash bargaining over T_i: they apportion the net gain of agent i after he has compensated agent j. Why do they do so? Simply because it would be unfair to let agent i get all the relative gain after the switch to unilateral communication: in fact, it is reasonable to assume that, if agents can somehow increase the total sum of their individual profits, they both have a role in doing it. Intuitively, an agent's share should increase in her outside option, and decrease in the other agent's one. The use of a Nash bargaining procedure can solve the problem.

The compensation mechanism ensures that the social gains from unilateral communication are fairly redistributed among agents. As we have presented it, compensation mechanism can be graphically explained in Figure 1.10, where *ex ante* expected profits increase from (a) to (d).

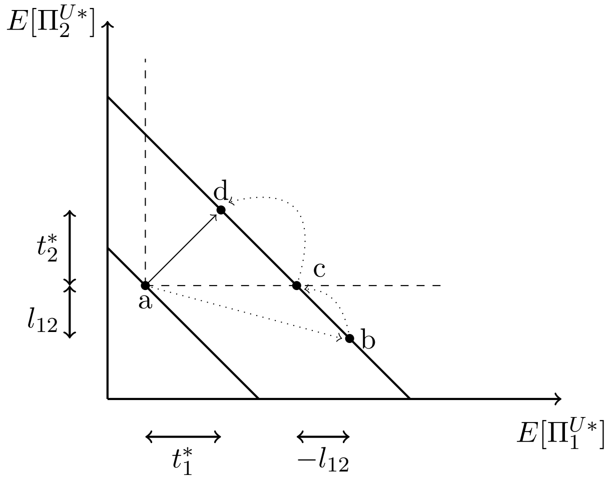


FIGURE 1.10: A possible compensation mechanism.

Imagine that, when they do not interact with each other, agents 1 and 2 earn expected profits such as they were in point (a): in this point, agent 2 is earning greater expected profits of agent 1. Suppose they decide to start unilateral communication where agent 1 "talks" and agent 2 "listens". After the change of setting, not only agent 1 ends up earning more profits than agent 2, but agent 2 is even earning less profits than before, as in point (b). The first step of the compensation mechanism s_{12} is thus directed from agent 1 to agent 2 and it makes agent 2 indifferent between no interaction and unilateral communication, as in point (c). The second step from point (c) to point (d) is the outcome of Nash bargaining over the extra profits earned by agent 1 in the new setting, where the share t_1^* goes to agent 1 while the share t_2^* goes to agent 2. In this bargaining problem, agents split the social gain according to the following optimization problem:

$$\begin{aligned} \max_{t_1,t_2} \quad & \big(t_1 - E_1\big[\Pi_1^*(\alpha_1,\alpha_2)\big]\big)\big(t_2 - E_2\big[\Pi_2^*(\alpha_1,\alpha_2)\big]\big) \\ \text{s.t.} \quad & t_1 + t_2 = E_1\big[\Pi_1^{U*}(\alpha_1,\alpha_2)\big] + E_2\big[\Pi_2^{U*}(\alpha_1,\alpha_2)\big] - E_1\big[\Pi_1^*(\alpha_1,\alpha_2)\big] - E_2\big[\Pi_2^*(\alpha_1,\alpha_2)\big] \end{aligned}$$

The solution is the following:

$$(t_1^*, t_2^*) = \left(\frac{E_1\big[\Pi_1^{U*}\big] + E_2\big[\Pi_2^{U*}\big]}{2} - E_2\big[\Pi_2^*\big], \frac{E_1\big[\Pi_1^{U*}\big] + E_2\big[\Pi_2^{U*}\big]}{2} - E_1\big[\Pi_1^*\big]\right)$$

where we omitted the notation (α_1, α_2) due to space limits.

As expected, the share of agent i decreases in the outside option of agent j and hence increases in the outside option of agent i: i.e. unilateral communication should reward more who was in a "better position" before starting the communication. □

Proposition 1.5. *Both agents have incentives to send untruthful messages.*

Proof. Recall that the only way in which a generic agent i can rationally lie is issuing $M_i = L$. Moreover, recall that untruthful messages are credible only when they are plausible.

We first prove agent 1 has incentives to send untruthful messages. The message $M_1 = L$ is plausible only when $S_2^a = S_2^L$. In this case, if agent 2 trusts agent 1's lie, agent 2 would produce $x_2^{L,L*}$ instead of

$x_2^{H,L*}$, and the latter quantity is always smaller than the former. Because of this, agent 1 can defect and increase the production above the truthful level, gaining additional profits. Hence, as long as $\alpha_1 \leqslant \alpha_2$, there is a nonnegative probability that sending untruthful messages is profitable for agent 1, so he has incentives to lie.

We now prove that agent 2 has incentives to lie, sending M_2^L. When $S_2^a = S_2^L$, she is actually telling the truth, but when $S_2^a = S_2^H$ she is sending an untruthful and plausible message. In this latter situation, if she communicates M_2^L and agent 1 trusts her, agent 1 would produce $x_1^{H,L*}$, which is a smaller quantity with respect to $x_1^{H,H*}$ he would have produced if agent 2 told the truth. Because of this, agent 2 can defect and increase the production above the truthful level, gaining additional profits. □

Appendix B

Homogeneity in categorical thinking

In this section, we examine a special case of the previous analysis, i.e. that of homogeneity in categorical thinking. Agents with homogeneous categorical thinking share the same threshold: in other words, $\alpha_1 = \alpha_2 = \alpha$. This may occur because they have developed similar cultural traits and hence similar ways to interpret available information.

Homogeneous agents are exchangeable, in the sense that given the same optimization problem, they come up with the same solution. They do it *a priori*: it is not a coordination game, nor they are trying to mimic each other; they *behave* in the same way because they *think* in the same way.

Imposing $\alpha_1 = \alpha_2$ has an important consequence: communication ceases to exist, not because it is not truthful but because it is useless. The categories agents observe not only "have the same label" (both agents observe S_i^L or S_i^H alternatively), but are also identical: for $k \in \{L, H\}$, if $\alpha_1 = \alpha_2$ then $S_i^k = S_j^k$. All possible messages are predictable, and agents would simply avoid to send them.

In this setting, the *interim* expected payoff for a representative agent i are the same as in (1.5), but now we have

$$E_i\left[x_j|S_i^k\right] = x_i$$

for $k \in \{L, H\}$. The maximization of *interim* expected profits lead to two optimal production quantities, one for each category observed:

$$x_i^{k*} = \frac{E_i\left[A|S_i^k\right]}{3}$$

for $i = 1, 2$, and where $E_i\left[A|S_i^L\right] = \frac{\alpha}{2}$ and $E_i\left[A|S_i^H\right] = \frac{1+\alpha}{2}$. *Interim* expected payoffs are

$$E_i\left[\Pi_i^*|S_i^L\right] = \frac{\alpha^2}{36}$$

for the L situation, and

$$E_i\left[\Pi_i^*|S_i^H\right] = \frac{(1+\alpha)^2}{36}$$

for the H situation. *Ex ante* expected payoffs are

$$E_i\left[\Pi_i^*(\alpha)\right] = \frac{1}{36}\left(1 + \alpha - \alpha^2\right) \tag{1.11}$$

Ex ante expected payoffs, which only depend on α, have an inverted u-shape with vertex in $\alpha = \frac{1}{2}$, implying that $\alpha_1 = \alpha_2 = \frac{1}{2}$ are indeed the thresholds that maximize *ex ante* expected profits when agents have homogeneous categorical thinking. With these thresholds, agents divide the support in two equally sized categories: low and high categories are equally probable, and agents' production decisions are, on average, the best ones.

Nonetheless, a comparison between the *ex ante* expected payoffs computed for agents with heterogeneous thresholds and those for agents with homogeneous thresholds in shows that agents with different thresholds are better off than agents with equal thresholds, if the first ones have complementary thresholds (i.e. $\alpha_1 + \alpha_2 \approx 1$). In other words, we have

$$E_i\left[\Pi_i^*(\alpha_1 \neq \alpha_2)\right] > E_i\left[\Pi_i^*(\alpha_1 = \alpha_2)\right] \tag{1.12}$$

for each $i = 1, 2$ if $\alpha_1 + \alpha_2 \approx 1$.

Having the same threshold hampers the achievable profits: this is due to increased competition within each situation. Having the same threshold subtly implies that agents face, with certain probabilities, two different situations, and in each one they behave in the same way. Instead, when agents have different thresholds, each of them is more informed than the other one in one segment of the market: in particular, agent 1 has a more precise information when he observes S_1^L than agent 2 has when she observes S_2^L. This gives him an advantage when the market size is indeed low. On the contrary, this advantage disappears and goes in favor of agent 2 when the market size is big. Overall, this information advantage creates positive spillovers, as it allows agents to expect higher payoff compared to the situation in which they had no advantage, i.e. when agents have equal threshold.

Bibliography

[Allport, 1954] Allport, G.W., 1954. *The Nature of Prejudice*, Cambridge (MA): Addison–Wesley.

[Bacharach, 2003] Bacharach, M., 2003. "Framing and cognition in economics", in Dimitri, N., Basili, M. and Gilboa, I. (eds.) *Cognitive Processes and Economic Behavior*, London: Routledge, 63–74.

[Baumol and Quandt, 1964] Baumol, W.J. and Quandt, R.E., 1964. "Rules of thumb and optimally imperfect decisions", *The American Economic Review*, **54** (2), 23–46.

[Bednar and Page, 2007] Bednar, J. and Page, S., 2007. "Can games(s) theory explain culture? The emergence of cultural behavior within multiple games", *Rationality and Society*, **19** (1), 65–97.

[Benhabib et al., 2011] Benhabib, J., Jackson, M.O. and Bisin, A., 2011. *Handbook of Social Economics*, Amsterdam: Elsevier.

[Cohen and Lefebvre, 2005] Cohen, H. and Lefebvre, C., 2005. *Handbook of Categorization in Cognitive Science*, Amsterdam: Elsevier.

[Denzau and North, 1994] Denzau, A. and North, D., 1994. "Shared mental models: Ideologies and institutions", *Kyklos*, 47 (1), 3–31.

[Fryer and Jackson, 2008] Fryer, R. and Jackson, M.O., 2008. "A categorical model of cognition and biased decision making", *The B.E. Journal of Theoretical Economics*, **8** (1, Contributions), Art. 6.

[Gibbons et al., 2017] Gibbons, R., LiCalzi, M., and Warglien, M., 2017. "What situation is this? Coarse cognition and behavior over a space of games", *Working papers 09*, Department of Management, Ca' Foscari University of Venice.

[Halevy et al., 2012] Halevy, N., Chou, E.Y., and Murnighan, J.K., 2012. "Mind games: The mental representation of conflict", *Journal of Personality and Social Psychology*, **102** (1), 132–148.

[Heller and Winter, 2016] Heller, Y. and Winter, E., 2016. "Rule rationality", *International Economic Review*, **57** (3), 997–1026.

[Hoff and Stiglitz, 2010] Hoff, K. and Stiglitz, J.E., 2010. "Equilibrium fictions: A cognitive approach to societal rigidity", *American Economic Review*, **100** (2), 141–146.

[Hoff and Stiglitz, 2016] Hoff, K. and Stiglitz, J.E., 2016. "Striving for balance in economics: Towards a theory of the social determination of behavior", *Journal of Economic Behavior & Organization*, **126** (B), 25–57.

[LiCalzi and Mühlenbernd, 2019] LiCalzi, M. and Mühlenbernd, R., 2019. "Categorization and cooperation across games", *Games*, **10** (1), 5.

[Mengel, 2012] Mengel, F., 2012. “Learning across games”, *Games and Economic Behavior*, **74** (2), 601–619.

[Mengel, 2012b] Mengel, F., 2012. “On the evolution of coarse categories”, *Journal of Theoretical Biology*, **307**, 117–124.

[Mohlin, 2014] Mohlin, E., 2014. “Optimal categorization”, *Journal of Economic Theory*, **152**, 356–381.

[Mullainathan et al., 2008] Mullainathan, S., Schwartzstein, J., and Shleifer, A., 2008. “Coarse thinking and persuasion”, *The Quarterly Journal of Economics*, **123** (1), 577–619.

[Rovelli, 2017] Rovelli, C., 2017. *L’Ordine del Tempo*, Milano: Adelphi.

[Samuelson, 2001] Samuelson, L., 2001. “Analogies, adaption, and anomalies”, *Journal of Economic Theory*, **97** (2), 320–366.

[Tanaka and Taylor, 1991] Tanaka, J.W. and Taylor, M., 1991. “Object categories and expertise: Is the basic level in the eye of the beholder?”, *Cognitive Psychology*, **23** (3), 457–482.

Chapter 2

Tacit vs Explicit Collusion Between Differently Informed Firms

Abstract

Collusion between firms is not always illegal: it is the act of communicating to coordinate behavior that makes it illicit. In this paper, we analyze the incentive to start and maintain tacit (legal) and explicit (illegal) collusion of two firms that compete in a Cournot duopoly, producing perfect substitutes for a mass of consumers. We assume that firms are differently informed about the market demand, and we prove that tacit collusion is easier to sustain in the long run, because firms have little incentives to share their private information and enter an explicit collusion, which might be more profitable although it requires side payments. This result has interesting welfare implications. While consumers are generally worse off when producers collude, we identify a precise condition under which Competition Authorities could mitigate these negative effects: namely, centralizing the provision of information about market demand.

Keywords: Cournot duopoly, Tacit collusion, Explicit collusion, Signals.

JEL Classification Numbers: D43, D83, L13.

2.1 Introduction

In 1890, the US Sherman Act shed light on the prominent problem of anti-competitive agreements among firms. To help overcome these problems, Competition Authorities (CAs) were established in several countries throughout the twentieth century. Since their introduction, the main objective of CAs, such as the ICA (or AGCM) in Italy or the FTC in the US (whose positive effective impact has recently been investigated by [Bos et al., 2018]), has been the support of competitive markets and the prevention of market behaviors apt to harm consumers' welfare. Anti-competitive behaviors such as price fixing, market sharing, and abuses of dominant positions are discouraged and punished. In fact, such conducts may not only decrease consumers' welfare, but even the total welfare of the society.[1] The importance of the topic regarding anti-competitive concertations and their formation have grown throughout the years, and the literature has investigated which collusive agreements are more likely to emerge and the factors that may facilitate (or hinder) these illegal behaviors. However, there are still open questions: among them, the role of asymmetries in firms' perception of the market demand and the consequent attainability of different types of collusion.

In this chapter, we study an infinitely repeated Cournot game by two firms with asymmetric and private information about the market demand. We investigate the effects of these asymmetries on firms' incentives to collude, either tacitly or explicitly (the main difference is that in the latter there is exchange of communication regarding market demand, whereas in the former there is not), and the conditions under which firms can reach and maintain such cooperation agreements. In particular, we concentrate on firms' discount rate, the critical parameter that determines the effectiveness of collusive intertemporal incentive schemes. We find that minimum discount rates that consent collusive practices are correlated with the (imperfectly observed) market demand: the higher the market demand, the higher the minimum discount factor needed to attain collusion. The main finding of this chapter is that explicit (or overt) collusion is more difficult to sustain with respect to tacit collusion: the minimum discount factors needed to attain explicit collusion is always higher than the minimum discount factor needed to attain tacit collusion. Hence, only very patient firms would explicitly collude in the long-run: this is mainly due to incentives to misreport communication in order to deceive the opponent. Conversely, even relatively less patient firms may decide to tacitly collude.

According to the international market regulation, not all forms of collusion are illegal: only explicit collusion is deemed to be illicit, whereas tacit collusion, despite leading to economic outcomes similar to the previous one, is not considered illegal.[2] The main difference between them is the presence of communication between the parties involved: "it is the act of communicating to coordinate behavior that is illegal (or taken as evidence of illegality), and not the actual prices that are charged" ([Bos et al., 2018]: 377). The finding that tacit collusion can be sustained even when explicit collusion cannot is potentially bad news for Competition Authorities: although the latter can be discouraged with leniency programs ([Spagnolo, 2008]) and accurate investigations, the former is legal and difficult to trace. For instance, in the last few years (along with the boom of the online sales and e-commerce) there has been an increase in the adoption of algorithms that may help different firms that operate in the same market collude tacitly, i.e. without the use of

[1] Under certain conditions, the marginal increase in the producers' surplus is smaller than the marginal decrease in consumers' surplus: what is lost is usually called "deadweight loss".

[2] "Any information exchange with the objective of restricting competition on the market will be considered as a restriction of competition by object" (§72, European Commission Giudelines on the applicability of Article 101 TFUE).

direct or indirect communication. Tacit collusion may be fostered by algorithms, as they enlarge the transparency of the market and have the ability to detect small price changes and react accordingly (see, for instance, [Ezrachi and Stucke, 2016], [Deng, 2018], and [Calvano et al., 2019]).

In the final part of this chapter we focus on a welfare analysis, looking at a possible method through which a Competition Authority might increase consumers' surplus: centralize the provision of market information and increase the information level of the firms.

The chapter is organized as follows. The rest of this section assesses the related literature. Section 3.2 lays down the model, the assumptions and the static equilibrium. Sections 2.3 and 2.4 analyze, in a dynamic setting, tacit and explicit collusion, respectively. Results are stated in Section 2.5, welfare analysis in Section 2.6. Section 3.6 presents conclusions and possible extensions. Appendix A includes most of the formulas, which are too long to be reported in the main text, while all propositions and corollaries are proved in Appendix B.

2.1.1 Related literature

Collusion is one of the possible practices through which producers may increase their profits. As reported in [Levenstein and Suslow, 2006]: 45, "producers form cartels with the goal of limiting competition to increase profits. By restricting output and increasing price, ideally to the price a monopolist would set, profits are jointly maximized". As mentioned by [Motta, 2004], there are two main conditions that generate an incentive compatible structure that allows the existence of cartels or, more in general, collusion: (i) detectability of deviations, and (ii) the presence of credible punishments. The first condition holds when all participants are able to "*detect* in a timely way that a deviation [...] has occurred" ([Motta, 2004]: 139), while the second one guarantees that any deviation is punished by the rivals, depressing the profits of the deviator. In other words, deviators must be first recognized, and then punished. Communication may be useful for potentially colluding firms: in fact, under explicit collusion firms can exchange information with each other and coordinate on their jointly preferred equilibrium without having to experiment with the market, which might be costly ([Motta, 2004]). In this sense, communication helps coordination, which may be difficult to reach when collusion is tacit ([Harrington, 2008]).

The literature has considered several factors that can facilitate (or hinder) collusion. Among those that facilitate it, there are a sufficiently limited number of firms in the market, industries' concentration, product homogeneity, and more in general symmetries between firms (see [Motta, 2004] and [Vives, 2009] for a more extensive inquiry). For instance, [Hay and Kelley, 1974] find that collusion is most likely to occur and endure when the number of firms is small, industry-concentration is high, and the product is homogeneous. [Levenstein and Suslow, 2006] examine a wide variety of empirical studies of cartels, finding that collusion can emerge even in unconcentrated industries (but almost always because firms rely on industry associations). [Green and Porter, 1984] assess the role of industry stability, proving that the evidence of competitive prices is not a sufficient condition to say that cartels are absent. Recently, all the factors that may facilitate collusion have been object of new analyses, deepening both the theoretical and the empirical discourse. For instance, [Garrod and Olczak, 2018] revisited the types of market structures in which cartels are more likely to arise, due to the "mounting evidence that some cartels do not seem to arise in markets that standard theory predicts" ([Garrod and Olczak, 2018]: 2).

In the literaure, there is a general consensus that firms' asymmetries hinder collusion. For instance, [Davies et al., 2011] empirically identify the market structures that a CA is likely to associate with (tacit) collusion: these are markets characterized by high concentration and a limited number of firms with symmetric market shares. [Compte et al., 2012] find mixed evidence. They explore the issue of asymmetric capacities in a Bertrand duopoly, showing that asymmetries make collusion more difficult to sustain when the aggregate capacity is limited, but not when the aggregate capacity is much larger than the market size. [Garrod and Olczak, 2017] prove that asymmetries in firms' size do hinder collusion, even under imperfect monitoring. [Obara and Zincenko, 2017] show that a collusive agreement can be sustained even when firms have different discount factors: if firms' average discount factor falls strictly above the critical threshold $\frac{n-1}{n}$, then collusion can be sustained with n firms. [Athey and Bagwell, 2001] explore collusion in a repeated private information model with publicly-observed prices, finding that, despite being prohibited from making side-payments, firms can still implement self-enforcing schemes to sustain collusion.

At the best of our knowledge, there are not many works that investigate the effects of asymmetrically informed firms on the attainability of different types of collusive outcomes. Furthermore, the distinction between tacit and explicit collusion is frequently ignored ([Harrington, 2008]). The closest work to ours is [Garrod and Olczak, 2018], who capture the incentives for explicit collusion when firms can alternatively collude tacitly. In their model, authors assume capacity-constrained firms that, at the beginning of every period, may be part of a cartel or not. If there exists a cartel, each firm may decide either to explicitly collude with the counterparts, sharing its private information (price), or to become a whistleblower informing the Competition Authority on the existence of the cartel. If the firm is not part of a cartel, it may still decide to tacitly collude with the other firms. Authors find conditions for the emergence of tacit and explicit collusion. One of the prominent ones is that if the fines set by the CA are high enough, tacit collusion is not only more profitable than explicit, but also the only reachable collusive practice. Along with this article, our work contributes to shed light on the importance of tacit collusion with respect to explicit collusion.

Other works are consistent with further minor findings stated along the chapter: for instance, the result that collusion is more difficult to attain when the demand is relatively high is supported by [Rotemberg and Saloner, 1986]; moreover, as in the experiments run by [Fonseca and Normann, 2012] and [Waichman et al., 2014], we inspect the role of explicit communication and show that, under certain conditions, it leads to higher profits.[3]

To better understand the framework we are analyzing, consider the following example. There are two rental car companies which sell same-quality services (for instance, they both rent similar small cars). In their headquarters, these two companies have to decide how many cars move from a subsidiary to another one in order to match the forecast demand. Suppose that in a specific airport (for instance, Venice VCE) these two companies, with one subsidiary each, constitute a duopolistic market. A few days before the beginning of each month, the headquarters decide how many cars to provide their two subsidiaries. They try to figure out the demand of cars for that month, making reasonable predictions: this process includes considering possible events in the Venice area (carnival, the Venice Film Festival and so on) which can affect the demand. Each firm predicts the

[3] However, these two experiments find mixed results regarding the efficient number of firms involved in the communication. [Waichman et al., 2014] finds that two firms collude better than three, while [Fonseca and Normann, 2012] find that communication with four or six participant is more effective than with just two.

market demand according to its beliefs: *ceteris paribus*, although one firm might forecast a high demand, the other one might expect a low one. Agents may decide to collude in order to gain additional profits, and they may also decide to do so after explicitly exchanging information about their forecasts. We analyze the differences in the perception of the market demand and how they influence firms' incentives to collude, either tacitly or explicitly.

2.2 The model

Consider a Cournot duopoly played repeatedly for an infinite number of t periods. Firm i's only decision variable is the output quantity x_i (perfectly observable by both firms), and the cost function is assumed to be quadratic: $C(x_i) = cx_i^2$ with $c \in (0, 1)$. We assume the same c parameter (as if firms share the same production technology) and they both discount the future at the same discount rate $\delta \in (0, 1)$. The industry has aggregate output $x_i + x_j$ and it is characterized by the following inverse linear demand:

$$d = a - (x_i + x_j)$$

where a is the value of the market size. Profits of firm i are

$$\pi_i = \max\{x_i d - cx_i^2, 0\}$$

For simplicity, we let agent i embody firm i, without any substantial difference between these two labels.

Information asymmetry. The setting of this model resembles the one presented in Chapter 1. We assume that the market size is a uniformly distributed random variable A and each period t nature draws independent realizations a. We let $R_A = (0, 1)$ be the totally ordered support of A. Each agent i partitions R_A in two categories, which we denote C_i^L and C_i^H (as reminders for *Low* and *High*), such that $C_i^L \cap C_i^H = \emptyset$ and $C_i^L \cup C_i^H = R_A$. We call *threshold* $\alpha_i \in [0, 1]$ the personal parameter that separates two contiguous categories, hence we have that $C_i^L = (0, \alpha_i)$ and $C_i^H = (\alpha_i, 1)$. Without loss of generality, we assume $\alpha_1 \leqslant \alpha_2$. Figure 2.1 is a possible representation of how agents perceive the market size.

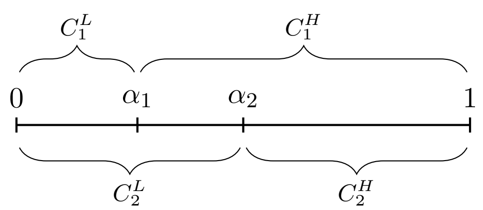


FIGURE 2.1: Example of agents' heterogeneous categorization of R_A.

We assume thresholds are common knowledge, so that agents know both their own categorization and the one of the competitor.

Signals. In Chapter 1 we assumed that at the beginning of the game (after a is realized) each agent perceives only the category in which a falls, and we have called this category *situation.* Here, we interpret this concept slightly differently. In fact, we assume that every period each agent receives a private signal that helps her realize whether the market size is *Low* or *High.* Perceived situations and observed signals are comparable concepts, and mathematically they can be treated in the same way. We let S_i^a denote the *signal* observed by agent i once a is realized. The signal S_i^a coincides with the category of C_i that includes a: for instance, if $a \in C_i^L$, then $S_i^a = S_i^L = C_i^L$. More in general, we have that $S_i^a = S_i^k = C_i^k$ if $a \in C_i^k$, with $i = 1, 2$ and $k \in \{L, H\}$. Because signal S_i^a depends on threshold α_i, each agent has an idiosyncratic interpretation of her private signal.

The signal S_i^a assists agent i in the decision making problem; every time she has to decide the production quantity, she can infer from the signal whether the market demand is *Low* or *High.*[4] In fact, after she has observed the signal, she refines her understanding of A updating her expectation $E_i[A|S_i^a]$ as a function of α_i.

Signals are endowed with accuracy, which depends on the categorization of the agent. Consider, for instance, an agent i with a threshold very close to 1, e.g. $\alpha_i = 0.9$. With 90% probability, agent i receives a *Low* signal, which is not very informative: she simply infers that the average market size is going to be $E_i[A|S_i^L] = 0.45$, but there are great chances that a falls relatively distant from this value. However, there is a 10% probability that she gets the *High* signal, which is very informative: there is almost null dispersion, so the production decision fits (almost) exactly the true market size. The trade-off between the accuracy and the probability of signals is simplified in Figure 2.2, which reports four examples: the signal is more accurate if the true value a falls in the smaller category rather then when it falls in the larger category.

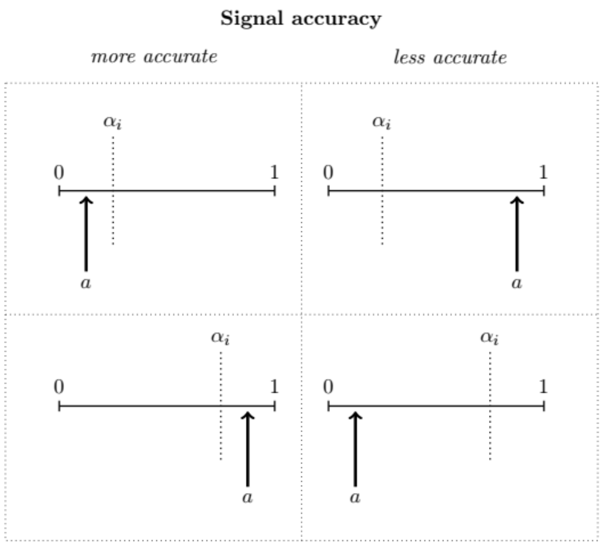


FIGURE 2.2: Signal accuracy

Collusion. As anticipated, we analyze two possible types of collusion: tacit and explicit. We define tacit collusion the non-verbal agreement between two firms to mutually decrease their production quantities in order to increase their profits: the agreement is *non-verbal* because no exchange of information regarding market demand is needed. The coordination

[4]Each signal is observed with a certain probability: $\Pr(S_i^a = S_i^L) = \alpha_i$ and $\Pr(S_i^a = S_i^H) = 1 - \alpha_i$, as A is uniformly distributed.

problem is solved by the decision to maximize the joint profit function, exactly as a single monopolist would do.

In explicit collusion, agents exchange relevant information about the market demand, learning from each other. They can send *messages* about the signal they have observed: let $M_i = \{M_i^L, M_i^H, \emptyset\}$ be the set of possible messages sent by agent i (to agent j). A message is an element of the set of categories of i: that is, if $M_i \neq \emptyset$, then $M_i \in C_i$. A message is truthful if it contains the true state of the world, i.e. if it coincides with the received signal, that is $M_i = S_i^a$. Agents are not forced to communicate, hence we include the possibility that they do not send any message, i.e. $M_i = \emptyset$ (exactly as in tacit collusion).

Collusion is Pareto-efficient but individually not a dominant strategy in a one-shot interaction: both agents have incentives to deviate, increasing their production quantities and obtaining a higher level of profits. In a repeated interaction, impatience is a driver: in each period, agents face the decision either to maintain the cartel or to renege on the collusion agreement, and they make this decision comparing the present discounted values of the utilities implied by these two alternatives.

We acknowledge the criticism of [Harrington, 2017], for which any kind of collusion would imply some sort of communication: in his view, "tacit" is simply a vacuous label. As it is used in the literature, we use it as well in order to discriminate between two different types of cartels that lead to different results for consumers and producers. In this sense, we remain silent on how the collusion agreement that we call "tacit" is reached: in our work, the world "communication" only regards an exchange of information about market demand.

When they set their production quantities, agents decide dependently on the market demand they expect. The expected market demand for agent i is

$$E_i\Big[D\Big|S_i^a, M_j\Big] = E_i\Big[A|S_i^a \cap M_j\Big] - \Big(x_i + E_i[x_j|S_i^a, M_j]\Big) \tag{2.1}$$

and depends on the signal about the market size received by agent i (S_i^a), the message (if there is any) received by the opponent (M_j), the expected quantity set by the opponent ($E_i\Big[x_j|S_i^a, M_j\Big]$), and the own quantity x_i. Using (2.1), we can write the *interim* expected profits for agent $i = 1, 2$ as

$$E_i\Big[\Pi_i\Big|S_i^a, M_j\Big] = \max\Big\{x_i E_i\Big[D\Big|S_i^a, M_j\Big] - cx_i^2, 0\Big\} \tag{2.2}$$

We assume agents are colluding (tacitly or explicitly). At the beginning of every period, agents decide whether to keep colluding or to renege on the collusive agreement. Once nature has drawn a and production decisions are made, *ex post* payoffs are realized and the true market size is observed at the end of the period. Besides impatience, collusion has to be consistent with precise actions and production decisions. We assume that firms can *a fortiori* infer (from the quantity produced by the opponent) whether the opponent colluded or not.

In both types of collusion, we assume the harshest possible punishment after defection: relentless Nash reversion to spot competition. Hence, we satisfy Motta's conditions for the existence of cartels: every deviation is detectable (this is guaranteed by observability of production quantities for tacit collusion and verifiability of information for explicit), and punishments exist and are credible.

Before analyzing the two types of collusion, we compute the static equilibrium.

2.2.1 Static equilibrium

In spot competition (and with no communication exchange), each agent maximizes her *interim* profits $E_i\left[\Pi_i \middle| S_i^a\right]$ given the signal S_i^a. Agent i does not know with certainty whether agent j has observed a low or high signal; for instance, when $S_1^a = S_1^H$ then S_2^a could correspond to S_2^L with a certain probability or to S_2^H with the complementary probability. Nonetheless, in certain cases agents can infer with certainty the signal observed by the opponent; for instance, if $S_1^a = S_1^L$ then $S_2^a = S_2^L$, and if $S_2^a = S_2^H$ then $S_1^a = S_1^H$.[5]

Assuming nonnegative profits, we can use *interim* profits in (2.2), where $M_j = \emptyset$, to describe *ex ante* profits:

$$E_i\Big[\Pi_i\Big] = \Pr(S_i^a = S_i^L)E_i\Big[\Pi_i|S_i^L\Big] + \Pr(S_i^a = S_i^H)E_i\Big[\Pi_i|S_i^H\Big] \tag{2.3}$$

Each agent $i = 1, 2$ sets her production quantity in order to maximize her *interim* profits, i.e. each agent solves the following maximization problem"

$$\max_{x_i^k} \; E_i\Big[\Pi_i|S_i^k\Big] = x_i^k\Big(E_i\Big[A|S_i^k\Big] - (x_i^k + E_i\Big[x_j|S_i^k\Big])\Big) - C(x_i^k) \tag{2.4}$$

for each $k \in \{L, H\}$, where $E_i[x_j|S_i^a]$ is a convex combination of x_j^L and x_j^H, computed as $E_i\Big[x_j|S_i^a\Big] = \Pr\Big(S_j^a = S_j^L|S_i^a\Big)x_j^L + \Pr\Big(S_j^a = S_j^H|S_i^a\Big)x_j^H$.

Solving (2.4) we can compute the equilibrium production quantities, that we call $x_1^{L,spot}$, $x_1^{H,spot}$, $x_2^{L,spot}$ and $x_2^{H,spot}$. These quantities are used to compute the equilibrium *interim* profits, that we denote $E_i[\Pi_i^{spot}|S_i^a]$, and the equilibrium *ex ante* profits $E_i[\Pi_i^{spot}]$. *Ex ante* profits $E_i[\Pi_i^{spot}]$ increase as threshold α_i approaches 0.5 and threshold α_j approaches its boundary value (either 0 or 1). For instance, $E_1[\Pi_1^{spot}]$ is maximum when $\alpha_1 = 0.5$ and $\alpha_2 = 1$, while $E_2[\Pi_2^{spot}]$ is maximum when $\alpha_1 = 0$ and $\alpha_2 = 0.5$. In these points, in fact, we say that one agent is *best-informed*, whereas the other one is *uninformed*: the former can exactly divide the market demand in two equal segments (categories), while the latter can distinguish only a singular (coarse) category.

In the following section, we analyze agents' incentives to tacitly collude and create (and maintain) a cartel in a repeated interaction.

2.3 Tacit collusion

In this section, we assume firms are tacitly agreeing on forming a cartel. First, we analyze their production quantities and profits: when they collude, firms limit their productions, yielding a price increase and hence an increment in their expected profits. Then, we compute firms' minimum discount factors needed to attain tacit collusion. We find that minimum discount factors correlate with the market size: when firms observe a *high* market size, they need a higher discount factor in order to resist the temptation to renege on collusion.

Under efficient tacit cooperation, agents behave as if they were a single monopolist, and maximize the joint *ex ante* profits, which is the sum of their individual *ex ante* profits. They do so without exchanging any relevant information; i.e. $M_i = \emptyset$ for each $i = 1, 2$.

[5]We have that $\Pr(S_2^L|S_1^H) = \frac{\alpha_2-\alpha_1}{1-\alpha_1}$, $\Pr(S_2^H|S_1^H) = \frac{1-\alpha_2}{1-\alpha_1}$, $\Pr(S_1^L|S_2^L) = \frac{\alpha_1}{\alpha_2}$, and $\Pr(S_1^H|S_2^L) = \frac{1-\alpha_1}{\alpha_2}$.

The maximization problem becomes:

$$\max_{x_i^L, x_i^H, x_j^L, x_j^H} E\Big[\Pi_i + \Pi_j\Big]$$

From this maximization problem we can compute the optimal quantities, that we call $x_1^{L,tacit}$, $x_1^{H,tacit}$, $x_2^{L,tacit}$ and $x_2^{H,tacit}$. Using these quantities, we can derive the *interim* profit functions $E_i\Big[\Pi_i^{tacit}|S_i\Big]$, and the *ex ante* profits $E_i\Big[\Pi_i^{tacit}\Big]$ (weighted by the probability of each signal).

Tacit collusion is always beneficial for both agents at the same time; in fact, they both reduce their production quantities ($x_i^{k,tacit} < x_i^{k,spot}$), increasing their profits in every segment of the market ($E_i\Big[\Pi_i^{tacit}|S_i^k\Big] > E_i\Big[\Pi_i^{spot}|S_i^k\Big]$ for each $k \in \{L, H\}$).

When they are involved in a cartel, agents (independently of the signal received) have an incentive to deviate from the collusion agreement increasing their production and, hence, expanding their profits. We call $\tilde{x}_i^{k,tacit}$ the optimal quantity produced in defection by agent i after observing signal S_i^k knowing that agent j would produce the tacit collusion quantity. Call $E_i\Big[\tilde{\Pi}_i^{tacit}|S_i^k\Big]$ the expected profits of agent i after defection, given signal S_i^k. We have that

$$\tilde{x}_i^{k,tacit} > x_i^{k,spot} > x_i^{k,tacit}$$

and

$$E_i\Big[\tilde{\Pi}_i^{tacit}|S_i^k\Big] > E_i\Big[\Pi_i^{tacit}|S_i^k\Big] > E_i\Big[\Pi_i^{spot}|S_i^k\Big]$$

for each $k \in \{L, H\}$.

In a one-shot interaction, tacit collusion is not an equilibrium: the incentives to renege are too large. However, in a repeated game we can study the minimum discount factor needed to attain collusion in the long run.

At the beginning of any period t, both agents privately decide whether to renew the cartel or to defect. Reneging implies two main consequences: a (relatively) high profit at the time t of defection, but a (relatively) smaller profit for the rest of the game, i.e. from $t+1$ onward. In fact, Nash reversion and information verifiability ensure that after any defection, independently of the signal received, the agent which suffers the deviation retaliates switching to spot competition forever. The incentive compatibility constraint for agent i can be generally written as follows:

$$\Big(1-\delta\Big)\left(E_i\Big[\Pi_i^{tacit}|S_i^k\Big] + \sum_{t=1}^{\infty}\delta^t E_i\Big[\Pi_i^{tacit}\Big]\right) \geqslant (1-\delta)\left(E_i\Big[\tilde{\Pi}_i^{tacit}|S_i^k\Big] + \sum_{t=1}^{\infty}\delta^t E_i\Big[\Pi_i^{spot}\Big]\right)$$

and rewritten as

$$\delta \geqslant \frac{E_i\Big[\tilde{\Pi}_i^{tacit}|S_i^k\Big] - E_i\Big[\Pi_i^{tacit}|S_i^k\Big]}{E_i\Big[\Pi_i^{tacit}\Big] - E_i\Big[\Pi_i^{spot}\Big] + E_i\Big[\tilde{\Pi}_i^{tacit}|S_i^k\Big] - E_i\Big[\Pi_i^{tacit}|S_i^k\Big]} = \delta_i^{k,tacit} \tag{2.5}$$

for $i = 1, 2$ and $k \in \{L, H\}$. Call $\delta_i^{k,tacit}$ the RHS of inequality (2.5): it represents the minimum discount factor δ needed for agent i to decide to collude after she has received the signal $S_i^a = S_i^k$. Hence, there are in total four minimum discount factors (2 agents × 2 signals), but only two of them are binding. In fact, the minimum discount factor needed to attain collusion increases in the market demand; in other words, for a generic agent i the minimum δ is higher when $S_i^a = S_i^H$: $\delta_i^{H,tacit} > \delta_i^{L,tacit}$ for all $\alpha_1, \alpha_2 \in [0,1]^2$ (with $\alpha_1 \leqslant \alpha_2$) and $c \in (0,1)$. Exactly as found in [Rotemberg and Saloner, 1986] and explained by [Harrington, 2017], collusion is easier to maintain (i.e. it is more stable) when the margin for profit is sufficiently low. When market demand is big, and hence profits are high, the increase in current profits from reneging is high, while the continuation payoff (if neither of them reneges) is unaffected. This is summed up in the following Proposition 2.1:

Proposition 2.1. *In tacit collusion, $\delta_i^{H,tacit} = \max\left\{\delta_i^{H,tacit}, \delta_i^{L,tacit}\right\}$ represents the highest minimum discount factor, and hence the only one which is binding, for each agent $i = 1, 2$.*

Proof. See Appendix B. □

Given Proposition 2.1, we have that tacit collusion is self-enforcing as long as agents' common minimum discount factor δ is greater or equal to the highest of the two binding constraints $\delta_1^{H,tacit}$ and $\delta_2^{H,tacit}$. We let δ^{tacit} denote the highest of these two constraints. Analytically, tacit collusion is self-enforcing if

$$\delta \geqslant \delta^{tacit} = \max\left\{\delta_1^{H,tacit}, \delta_2^{H,tacit}\right\} \tag{2.6}$$

and hence δ^{tacit} is the minimum discount factor needed to attain tacit collusion. Clearly, δ^{tacit} depends on three parameters (α_1, α_2, and c), and the ranking between $\delta_1^{H,tacit}$ and $\delta_2^{H,tacit}$ also depends on the value of these parameters.

2.4 Explicit collusion

In this section, we assume firms are agreeing on forming a cartel with the help of explicit communication: when they collude explicitly, firms use communication to sustain their agreement. We analyze firms' production quantities under explicit collusion: after coordinating their actions, firms decide to limit their productions and increase their profits. Then, we investigate the minimum discount factors needed to sustain this kind of cooperation. Consistently with the previous section, we find that minimum discount factors correlates with the market size: the larger the market size, the higher the minimum discount factor needed for explicit collusion.

In explicit collusion, communication represents the mutual exchange of agents' private signals: agent i, after having observed her private signal S_i^a, may share it by sending a message M_i to agent j; the message is truthful only if $S_i^a = M_i$. Truthful communication is fundamental for explicit collusion. In fact, if one of the firms fools the other one reneging on the agreement in period t, the other one would notice it immediately at the end of the period, and switch to spot competition forever starting from period $t + 1$. Hence, we assume firms naively believe the messages they receive.[6]

When they exchange truthful messages, agents homogenize their expectations about the market demand: communication wipes out their interpretative differences. In fact, for each combination $S_1^k \cap S_2^q$ (where $k, q \in \{L, H\}$) agents have the same expectation $E_1[A|S_1^k \cap S_2^q] = E_2[A|S_1^k \cap S_2^q]$.

Communication increases the number of intervals of the market size perceived by each agent. To lighten notation, let S denote the agents' joint interpretation of the market size

[6] Consequently, receiving no message at all would be considered as a defection from the collusive agreement.

when they truthfully exchange signals: in other words, S is the combination of different agents' signals. Let $S \in \{L, M, H\}$ be the possible values of S, and each of these values identifies a combinations of signals S_1^a and S_2^a, such that

$$S = \begin{cases} L \text{ if } S_1^a = S_1^L \text{ and } S_2^a = S_2^L \\ M \text{ if } S_1^a = S_1^H \text{ and } S_2^a = S_2^L \\ H \text{ if } S_1^a = S_1^H \text{ and } S_2^a = S_2^H \end{cases}$$

Hence, we have that

$$\begin{aligned} E_i\Big[A|S = L\Big] &= E_i\Big[A|S_1^L \cap S_2^L\Big] \\ E_i\Big[A|S = M\Big] &= E_i\Big[A|S_1^H \cap S_2^L\Big] \\ E_i\Big[A|S = H\Big] &= E_i\Big[A|S_1^H \cap S_2^H\Big] \end{aligned}$$

for each $i = 1, 2$.

When communication is truthful, agents can team up and collude maximizing their joint *interim* profits, conditional on S. Hence, agent i' monopolistic maximization problem becomes the following:

$$\max_{x_i^k, x_j^k} \; E_i\Big[\Pi_i + \Pi_j|S = k\Big] = (x_i^k + x_j^k)\Big(E_i[A|S = k] - (x_i^k + x_j^k)\Big) - C(x_i^k) - C(x_j^k)$$

for each $k \in \{L, M, H\}$.

The problem leads to three systems of two best replies, for which the solution is represented by six pairwise-equal quantities. Call $x_i^{k,expl}$ the quantity produced by agent i under explicit collusion when $S = k$. Because agents have the same expectations about A, we have that $x_1^{k,expl} = x_2^{k,expl}$, and hence $E_1\Big[\Pi_1^{expl}|S = k\Big] = E_2\Big[\Pi_2^{expl}|S = k\Big]$ for each $k \in \{L, M, H\}$. *Ex ante* expected profits can be computed weighting *interim* profits for the probability of each combination S:

$$E_i\Big[\Pi_i^{expl}\Big] = \sum_k \Pr(S = k) \cdot E_i\Big[\Pi_i^{expl}|S = k\Big]$$

where $k \in \{L, M, H\}$.

In order to understand if explicit collusion is more profitable than spot competition, we compare the *ex ante* profits of each agent referred to these strategies. That is, agent i finds it profitable to truthfully bilaterally communicate (i.e. explicitly collude) if $E_i\Big[\Pi_i^{expl}\Big] \geqslant E_i\Big[\Pi_i^{spot}\Big]$. For agent 1, this condition becomes

$$E_1\Big[\Pi_1^{expl}\Big] \geqslant E_1\Big[\Pi_1^{spot}\Big] \tag{2.7}$$

and it is plotted in the left part of Figure 2.3. In grey, the locus of points for which condition (2.7) is satisfied: agent 1 finds it unprofitable to explicitly collude only when α_1 is sufficiently close to 0.5 and α_2 is sufficiently close to 1. These are the conditions under which agent 1 makes profits near their maximum value (which happens to be when $\alpha_1 = 0.5$ and $\alpha_2 = 1$). In this point, in fact, agent 2 is *uninformed*: she simply perceives a single segment of the market demand, and so agent 1 finds it advantageous to exploit this situation committing to spot competition. A similar argument can be applied to the following condition, valid for agent 2:

$$E_2\Big[\Pi_2^{expl}\Big] \geqslant E_2\Big[\Pi_2^{spot}\Big] \tag{2.8}$$

Condition (2.8) is plotted in the right part of Figure 2.3. These surfaces slightly increase or decrease according to cost c, so the following Figure 2.3 represents only an emblematic example (and it resembles the case when $c = 0.5$).

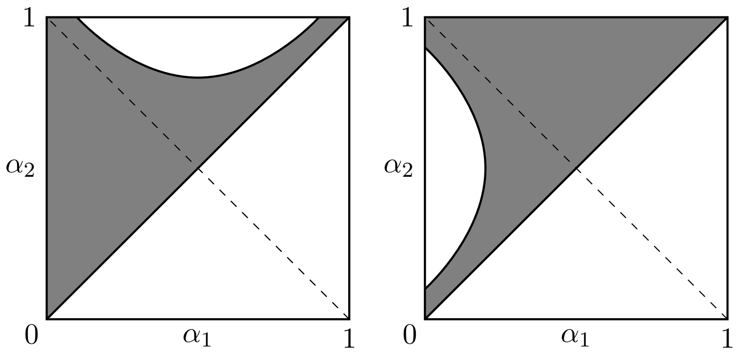


FIGURE 2.3: Left: area where agent 1 wants to explicitly collude. Right: area where agent 2 does.

If we superimpose the two graphs, we see that there exists conditions under which both agents *at the same time* find it profitable to explicitly collude: (i) when $\alpha_1 \approx \alpha_2$, and (ii) when $\alpha_1 + \alpha_2 \approx 1$, i.e. near the two diagonals. In the first case, communication is nearly useless: agents mostly gain from the fact that they are reducing their production quantities. In the second case, communication is fundamental: agents have *complementary* thresholds, and, if they share their interpretations, agent 1 is able to compensate for agent 2's lack of expertise regarding the low demand, whereas agent 2 compensate agent 1 when demand is high.

Side payments. There exist conditions under which explicit collusion cannot arise, because of the lack of incentives to share the private signal with a less-informed counterpart. Hence, we introduce a mechanism of side payments that encourages the more informed agent to share his information with the less-informed one, and that can sustain a mutual exchange of information.

First, we define who sells the information and who pays for it. Call *more informed* the agent with the threshold α_i closer to 0.5. For instance, if $\alpha_1 = 0.1$ and $\alpha_2 = 0.6$, agent 2 is *more informed*. When the distance from 0.5 is identical (i.e. when $\alpha_1 + \alpha_2 = 1$ or $\alpha_1 = \alpha_2$), we say that agents are *identically* (or *equally*) informed. Fees (whose amount is analyzed later) are paid by the *less informed* agent to the *more informed* one, only where they are needed, i.e. only in those areas for which there is no mutual advantage to collude explicitly. We assume that the payment is made before messages are exchanged (otherwise the payment should depend on the signal). The arrangement under which fees are paid is described in Figure 2.4, where we identify three regions (or areas) that we denote $r = W, N, C$ (in red, we have denoted the agent who has to pay the fee). The fee is paid by agent 1 in the white *west area* W, where agent 2 has no incentives to explicitly collude. Symmetrically, the fee is paid by agent 2 in the *north area* N, where it is agent 1 that has no incentives to collude explicitly. In the grey *central area* C, no fee is exchanged:

both agents do want to collude explicitly. Figure 2.4 represents a natural example of these areas.

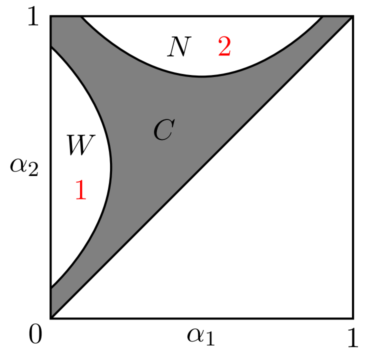


FIGURE 2.4: Graph of who pays the fee.

In the following analysis, we investigate agents' incentives to explicitly collude in each of these areas. Within every period t, actions develop as follows:

1. fee is paid by the *less informed* agent i to the *more informed* agent j (if $r = W, N$, as when $r = C$ no fee is exchanged);
2. messages are simultaneously exchanged;
3. production quantities are simultaneously set.

Before going on with the analysis, we show that agents have incentives to send untruthful messages. From Proposition 1.5 of Chapter 1, recall that the only way in which a generic agent $i = 1, 2$ can rationally lie is always issuing $M_i = M_i^L$. Moreover, recall that untruthful messages are credible only when they are plausible (i.e. the receiver cannot tell in advance if the message is truthful or not). The lie $M_1 = M_1^L$ issued by agent 1 is plausible only when $S_2^a = S_2^L$. In this case, if agent 2 trusts agent 1's lie, agent 2 would produce $x_2^{L,expl}$ instead of $x_2^{M,expl}$, and the latter quantity is always smaller than the former. Because of this, agent 1 can defect and increase the production above the truthful level, gaining additional profits. Hence, as long as $\alpha_1 \leqslant \alpha_2$, there is a nonnegative probability that sending untruthful messages is profitable for agent 1, so he has incentives to lie.[7] On the other hand, also agent 2 has incentives to lie, sending M_2^L (this lie is always plausible for agent 1). When $S_2^a = S_2^L$, she is actually telling the truth, but when $S_2^a = S_2^H$ she is sending an untruthful and plausible message. In this latter situation, if agent 1 trusts her he would produce $x_1^{M,expl}$, which is a smaller quantity with respect to $x_1^{H,expl}$ he would have produced if agent 2 told the truth. Because of this, agent 2 can defect and increase the production above the truthful level, gaining additional profits.

We analyze the west area of Figure 2.4 in Section 2.4.1; in Section 2.4.2, we analyze the north area. In Section 2.4.3, we analyze the central area.

2.4.1 Agent 1 pays the fee

In the west area $r = W$ of Figure 2.4, agent 1 pays the fee. Call β_1 the fixed share of agent 1's expected profits $E_1\left[\Pi_1^{expl}\right]$ that is transferred to agent 2 in order to make him

[7] There is also a positive probability that the lie is implausible: agent 2 realizes that agent 1 is lying. In this case, we assume that they will still collude for the current period, but switch to spot competition from the following period onward.

willing to collude explicitly. The share β_1 must be high enough to make agent 2 at least indifferent between explicit collusion and spot competition, and it must be low enough to make agent 1 not worse off than it would be in spot competition. Hence, we have two incentive compatibility constraints to satisfy; agent 1's incentive compatibility constraint is

$$(1-\beta_1)E_1\Big[\Pi_1^{expl}\Big] \geqslant E_1\Big[\Pi_1^{spot}\Big]$$

and it can be rewritten as

$$\beta_1 \leqslant \frac{E_1\Big[\Pi_1^{expl}\Big] - E_1\Big[\Pi_1^{spot}\Big]}{E_1\Big[\Pi_1^{expl}\Big]} = \bar{\beta}_1$$

where $\bar{\beta}_1$ is the upper bound of the share β_1. The ICC of agent 2 is

$$E_2\Big[\Pi_2^{expl}\Big] + \beta_1 E_1\Big[\Pi_1^{expl}\Big] \geqslant E_2\Big[\Pi_2^{spot}\Big]$$

which can be rewritten as

$$\beta_1 \geqslant \frac{E_2\Big[\Pi_2^{spot}\Big] - E_2\Big[\Pi_2^{expl}\Big]}{E_1\Big[\Pi_1^{expl}\Big]} = \underline{\beta}_1$$

where $\underline{\beta}_1$ is the lower bound of the fee, hence $\underline{\beta}_1 \leqslant \beta_1 \leqslant \bar{\beta}_1$. We use Nash bargaining to find the optimal share β_1^* given each agents' outside option. Hence, we solve the following problem:

$$\begin{aligned}
\max_{\beta_1,\beta_2} \quad & \left(E_2\Big[\Pi_2^{expl}\Big] + \beta_1 E_1\Big[\Pi_1^{expl}\Big] - E_2\Big[\Pi_2^{spot}\Big]\right)\left(\beta_2 E_1\Big[\Pi_1^{expl}\Big] - E_1\Big[\Pi_1^{spot}\Big]\right) \\
s.t. \quad & \beta_1 + \beta_2 = 1
\end{aligned}$$

where β_2 is the percentage of profits kept by agent 2. Solving the problem, we obtain that the optimal share β_1^* is

$$\beta_1^* = \frac{E_2\Big[\Pi_2^{spot}\Big] - E_2\Big[\Pi_2^{expl}\Big]}{E_1\Big[\Pi_1^{spot}\Big] + E_2\Big[\Pi_2^{spot}\Big] - E_2\Big[\Pi_2^{expl}\Big]}$$

Clearly, the share β_1^* decreases with respect to the outside option of agent 1 ($E_1\Big[\Pi_1^{spot}\Big]$). Figure 2.5 plots four examples of $\underline{\beta}_1$, $\bar{\beta}_1$ (dotted line) and β_1^* (solid line) for $\alpha_1 = [0, .05, .1, .15]$ from left to right, when $c = 0.5$. The examples show that, as we approach from left to right the end of the west area of Figure 2.4, the share $\beta_1^* \downarrow 0$: clearly, we have that $\beta_1^* \geqslant 0$ only if $r = W$.

The fee β_1^* is maximum (and equal to 10%, when $c = 0.5$) when $\alpha_1 = 0$ and $\alpha_2 = 0.5$, where the expected profits of agent 2 for spot competition are maximized: the compensation must be relatively high, and sufficient enough to clear out agent 2's incentives to switch to spot competition.

Is this mechanism sufficient to guarantee explicit collusion in the long run? We analyze, for each agent, the incentives to renege on the agreement.

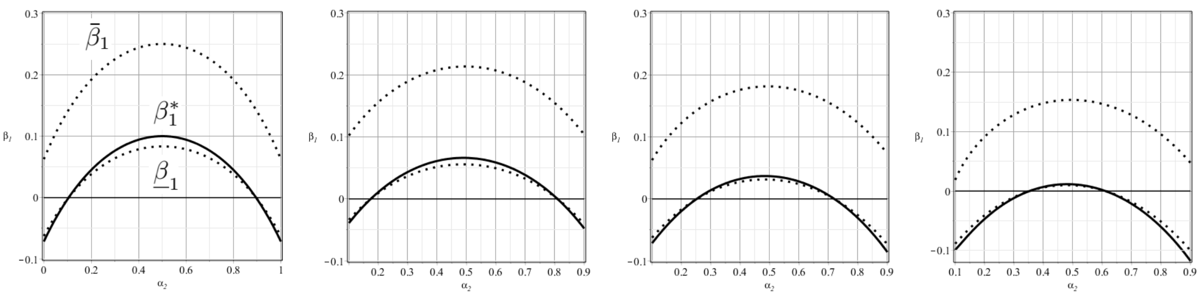


FIGURE 2.5: Upper bound, lower bound and optimal share for different value of α_1 (and $c = 0.5$).

Agent 1. At the beginning of each period, in consideration of the fact that he pays the fee up front, agent 1 can decide whether to (i) pay the fee, send a truthful message, and produce the collusion quantity, or (ii) pay the fee, renege on the agreement and fool agent 2 sending the (possibly) untruthful message M_1^L and producing a (slightly) bigger quantity (we rule out the possibility of not paying the fee: agent 2 would realize it immediately at time t and switch to spot competition, which harms agent 1).[8] If agent 1 decides to renege on the explicit collusion agreement, he produces a defection quantity $\tilde{x}_1^{k,expl}$ which is greater than the quantity that he is supposed to produce when colluding $x_1^{k,expl}$ for each $S = k$ and $k \in \{L, M, H\}$; moreover, $E_1\Big[\tilde{\Pi}_1^{expl}|S = k\Big] > E_1\Big[\Pi_1^{expl}|S = k\Big]$. This latter choice of reneging will be punished with Nash reversion. The following condition ensures that agent 1 will not deviate from the agreement:

$$\Big(1-\delta\Big)\bigg(E_1\Big[\Pi_1^{expl}|S=k\Big] - \beta_1^* E_1\Big[\Pi_1^{expl}\Big] + \sum_{t=1}^{\infty}\delta^t(1-\beta_1^*)E_1\Big[\Pi_1^{expl}\Big]\bigg) \geqslant$$
$$\geqslant (1-\delta)\bigg(E_1\Big[\tilde{\Pi}_1^{expl}|S=k\Big] - \beta_1^* E_1\Big[\Pi_1^{expl}\Big] + \sum_{t=1}^{\infty}\delta^t E_1\Big[\Pi_1^{spot}\Big]\bigg)$$

This can be rearranged as

$$\delta \geqslant \frac{E_1\Big[\tilde{\Pi}_1^{expl}|S=k\Big] - E_1\Big[\Pi_1^{expl}|S=k\Big]}{(1-\beta_1^*)E_1\Big[\Pi_1^{expl}\Big] - E_1\Big[\Pi_1^{spot}\Big] + E_1\Big[\tilde{\Pi}_1^{expl}|S=k\Big] - E_1\Big[\Pi_1^{expl}|S=k\Big]} = \delta_{1,W}^{k,expl} \quad (2.9)$$

For each $S = k$ with $k \in \{L, M, H\}$, call $\delta_{1,W}^{k,expl}$ the RHS of condition (2.9), where the subscript W indicates that we are dealing with the west area of Figure 2.4. Because $S \in \{L, M, H\}$, we have three possible values for $\delta_{1,W}^{k,expl}$. However, as we prove later, we can simplify the analysis stating that the minimum discount factor when $S = H$ is higher than any other case. We now analyze the incentives to deviate of agent 2.

Agent 2. Similarly to agent 1, at the beginning of every period t agent 2 faces a decision: (i) cash the fee, truthfully exchange messages and stick to the collusion production quantity, or (ii) cash the fee, (possibly) lie in the communication sending M_2^L and deceive agent 1 increasing the production accordingly, knowing that her deviation will be anyhow

[8]The strategy not to send any message is a dominated one: if firm i does not send a message to firm j at the beginning of period t, firm j would think that firm i is not willing to collude anymore, and will choose the spot competition quantities. The best reply for firm i would hence be to spot compete as well, without any additional benefit that, for instance, sending an untruthful message would carry.

punished by Nash reversion. Call $\tilde{x}_2^{k,expl}$ the optimal quantity produced by agent 2 after defection for each $S = k$ and $k \in \{L, M, H\}$. As agent 1, agent 2 can defect in three possible situations: however, clearly, she has more incentive to do so when $S = H$. We compute agent 2's minimum discount factor to attain collusion as

$$\Big(1-\delta\Big)\bigg(E_2\Big[\Pi_2^{expl}|S=k\Big] + \beta_1^* E_1\Big[\Pi_1^{expl}\Big] + \sum_{t=1}^{\infty} \delta^t \Big(E_2\Big[\Pi_2^{expl}\Big] + \beta_1^* E_1\Big[\Pi_1^{expl}\Big]\Big)\bigg) \geqslant$$

$$\geqslant (1-\delta)\bigg(E_2\Big[\tilde{\Pi}_2^{expl}|S=k\Big] + \beta_1^* E_1\Big[\Pi_1^{expl}\Big] + \sum_{t=1}^{\infty} \delta^t E_2\Big[\Pi_2^{spot}\Big]\bigg)$$

which can be rewritten as

$$\delta \geqslant \frac{E_2\Big[\tilde{\Pi}_2^{expl}|S=k\Big] - E_2\Big[\Pi_2^{expl}|S=k\Big]}{E_2\Big[\Pi_2^{expl}\Big] + \beta_1^* E_1\Big[\Pi_1^{expl}\Big] - E_2\Big[\Pi_2^{spot}\Big] + E_2\Big[\tilde{\Pi}_2^{expl}|S=k\Big] - E_2\Big[\Pi_2^{expl}|S=k\Big]} = \delta_{2,W}^{k,expl} \tag{2.10}$$

for each $S = k$ and $k \in \{L, M, H\}$. Call $\delta_{2,W}^{k,expl}$ the RHS of condition (2.10). Now that we have the minimum discount factors for both agents when agent 1 pays the fee (i.e. in $r = W$), we compute them for the remaining cases.

2.4.2 Agent 2 pays the fee

In the north area $r = N$ of Figure 2.4, agent 2 pays the fee. The following analysis is similar to the previous one: let β_2 denote the fixed share of agent 2's expected profits $E_2\Big[\Pi_2^{expl}\Big]$ paid to agent 1. As before, let $\underline{\beta}_2$ and $\bar{\beta}_2$ denote the incentive compatibility constraints of agent 1 and agent 2, respectively. The boundaries $\underline{\beta}_2$ and $\bar{\beta}_2$ are computed similarly (and symmetrically) to those in Section 2.4.1: we have that $\underline{\beta}_2 \leqslant \beta_2 \leqslant \bar{\beta}_2$ for all α_1, α_2 and c included in the north area of Figure 2.4. Solving the Nash bargaining problem, we find that

$$\beta_2^* = \frac{E_1\Big[\Pi_1^{spot}\Big] - E_1\Big[\Pi_1^{expl}\Big]}{E_1\Big[\Pi_1^{spot}\Big] + E_2\Big[\Pi_2^{spot}\Big] - E_1\Big[\Pi_1^{expl}\Big]}$$

Here, the fee β_2^* is maximum when $\alpha_1 = 0.5$ and $\alpha_2 = 1$, exactly the value for which agent 1 has maximum expected profits from the outside option (spot competition). The pattern of β_2^* resembles exactly the one of β_1^* depicted in Figure 2.5, but with α_1 in the x-axis.

We now compute the minimum discount factors needed to sustain explicit collusion when the fixed fee is paid by agent 2.

Agent 1. Being *more informed*, agent 1 cashes the fee at the beginning of each period and decides whether to stick to the agreement or renege on it increasing his production quantities and lie when it is possible. Agent 1 sticks to explicit collusion if, for $k \in$

$\{L, M, H\}$, the following condition holds:

$$\delta \geqslant \frac{E_1\left[\tilde{\Pi}_1^{expl}|S=k\right] - E_1\left[\Pi_1^{expl}|S=k\right]}{E_1\left[\Pi_1^{expl}\right] + \beta_2^* E_2\left[\Pi_2^{expl}\right] - E_1\left[\Pi_1^{spot}\right] + E_1\left[\tilde{\pi}_1^{expl}|S=k\right] - E_1\left[\Pi_1^{expl}|S=k\right]} = \delta_{1,N}^{k,expl} \tag{2.11}$$

Condition (2.11) is rearranged similarly to (2.9). We let $\delta_{1,N}^{s,expl}$ denote the RHS of (2.11), where the subscript N stands for "north".

Agent 2. After paying the fee, agent 2 could renege on collusion sending an untruthful message after she has observed a high signal and hence increase the production quantities above the collusion level. Agent 2 will honor the agreement if

$$\delta \geqslant \frac{E_2\left[\tilde{\Pi}_2^{expl}|S=k\right] - E_2\left[\Pi_2^{expl}|S=k\right]}{(1-\beta_2^*)E_2\left[\Pi_2^{expl}\right] - E_2\left[\Pi_2^{spot}\right] + E_2\left[\tilde{\Pi}_2^{expl}|S=k\right] - E_2\left[\Pi_2^{expl}|S=k\right]} = \delta_{2,N}^{k,expl} \tag{2.12}$$

for $k \in \{L, M, H\}$. Call $\delta_{2,N}^{k,expl}$ the RHS of (2.12).

2.4.3 No side payments

Lastly, we consider the central grey area $r = C$ of Figure 2.4, where no side payment is exchanged because both agents find it profitable to collude explicitly and exchange information rather than spot compete, even if (depending on their thresholds) one of them might be *more informed*. As in the previous cases, agents have incentives to deviate from the agreement. We now compute their minimum discount factors needed to attain explicit collusion, which, similarly to condition (2.5), can be generally written as

$$\delta \geqslant \frac{E_i\left[\tilde{\Pi}_i^{expl}|S=k\right] - E_i\left[\Pi_i^{expl}|S=k\right]}{E_i\left[\Pi_i^{expl}\right] - E_i\left[\Pi_i^{spot}\right] + E_i\left[\tilde{\Pi}_i^{expl}|S=k\right] - E_i\left[\Pi_i^{expl}|S=k\right]} = \delta_{i,C}^{k,expl} \tag{2.13}$$

The function $\delta_{i,C}^{k,expl}$ is the minimum discount factor for agent i needed to attain explicit collusion in the central grey area of Figure 2.4 when combination of signals $S = k$ is realized, where $k \in \{L, M, H\}$. In fact, both agents may decide to send an untruthful message and increase the production above the collusion level. Hence, we have six minimum discount factors (as in previous cases), but, for each agent i, only the highest one is binding.

Highest minimum discount factor. Computed the minimum discount factors for each area of Figure 2.4 under explicit collusion, we state the following Proposition 2.2.

Proposition 2.2. *In explicit collusion,* $\delta_{i,r}^{H,expl} = \max\left\{\delta_{i,r}^{H,expl}, \delta_{i,r}^{M,expl}, \delta_{i,r}^{L,expl}\right\}$ *represents the highest minimum discount factor for each agent* $i = 1, 2$ *and each region* $r = W, N, C$.

Proof. See Appendix B. □

This result is true for all $\alpha_1, \alpha_2 \in [0, 1]^2$ (with $\alpha_1 \leqslant \alpha_2$), for each $c \in (0, 1)$, and for each area $r = W, N, C$. Proposition 2.2 is coherent with Proposition 2.1. In fact, for both tacit and explicit collusion, minimum discount factors needed to attain collusion are correlated with the market size: the highest the (expected) market size, the highest the minimum discount rate.

Proposition 2.2 implies that, for each agent $i = 1, 2$, only one discount factor ($\delta_{i,r}^{H,expl}$ for $r = W, N, C$) is binding. Hence, when we want to understand if explicit collusion is feasible for the dyad, we only need to consider three pairs of two minimum discount factors ($\delta_{1,r}^{H,expl}$ and $\delta_{2,r}^{H,expl}$ for $r = W, N, C$). We can write the attainability condition for explicit collusion as follows:

$$\delta \geqslant \max\left\{\delta_{1,r}^{H,expl}, \delta_{2,r}^{H,expl}\right\} = \delta_r^{expl} \tag{2.14}$$

for each $r = W, N, C$. We denote by δ_r^{expl} the minimum discount factor needed to attain explicit collusion in each area r of Figure 2.4.

2.5 Results

In Sections 2.3 and 2.4 we have derived the highest minimum discount factors needed to sustain tacit and explicit collusion. In this section, we compare these discount factors and we find that those that regard explicit collusion are higher than those that regard tacit collusion. The main result of this chapter can be summarized in the following proposition:

Proposition 2.3. *Tacit collusion is easier to attain than explicit collusion: there are values of the common discount factor such that the former is an equilibrium while the latter is not.*

Proof. See Appendix B. □

This result is demonstrated by the following inequality:

$$\delta_r^{expl} \geqslant \delta^{tacit} \tag{2.15}$$

for each $r = W, N, C$. Inequality (2.15) is valid for all $\alpha_1, \alpha_2 \in [0, 1]^2$ (given $\alpha_1 \leqslant \alpha_2$) and $c \in (0, 1)$.

The main reason why explicit collusion is more difficult to attain is that it is linked to communication: in fact, agents have a strong incentive to misreport the privately observed signals and therefore lie when they exchange messages. Proposition 2.3 states that, when there are only two firms, illegal cartels are less likely to arise because tacit collusion is easier to attain. If we consider the discount factor as a cost, this result states that, in order to establish tacit collusion, firms face a smaller cost compared to explicit collusion. This result relates our work with the existing literature: a similar finding, in a different context, is identified by [Garrod and Olczak, 2018], for which, under certain conditions, tacit collusion is more appealing, rather then easier to attain.

This finding implies that, with respect to the work of Competition Authorities, tacit collusion represents a greater threat to natural competition than explicit collusion. In fact, the latter is not only easier to identify (because of “paper trails”) and discourage (for instance, through the establishment of leniency programs), but it is also more difficult for the firms to attain, because of the incentives to lie that we have discussed. On the contrary, tacit collusion is both (relatively) easy to attain even for less patient firms, and moreover it is difficult to detect and deter (by definition, tacit collusion leaves no trace and hence it is not illegal). Our result (despite being difficult to demonstrate empirically because of the very nature of tacit collusion) emphasizes that the distinction between explicit and tacit collusion, frequently neglected by the literature (as reported by [Harrington, 2008]) is instead significant, and it is relevant most notably for what concerns the *ex ante* conditions which can allow such collusive equilibria to emerge, rather than for the *ex post* effects, that are comparable overall. Moreover, tacit collusion might be particularly relevant in transparent and concentrated markets where the product is homogeneous, and where tacit collusion may be even fostered by the use of algorithms.[9]

Even though it is more difficult to reach, is explicit collusion also more remunerative than tacit collusion? The answer is not straightforward. In fact, the ranking between *ex ante* profits gained under explicit or tacit collusion depends on the values of α_1, α_2, and c. Explicit collusion, in fact, implies a trade-off that, under certain conditions, causes explicit collusion to be less profitable than tacit. This trade-off regards the level of competition that arises whenever firms exchange their private signals. In fact, on the one hand each firm may observe three possible combinations of signals about the market size instead of two private signals (as if they had the same ternary categorization): as stated by ([Athey and Bagwell, 2001]: 431), "the benefit from communication is that it allows firms to smoothly divide the market on a state-contingent basis". However, on the other hand the competition in each of these three intervals of the market size increases. This trade-off is positively solved (i.e. explicit collusion is more profitable than tacit) when agents are equally (or almost equally) informed with complementary thresholds, hence when $\alpha_1 + \alpha_2 \approx 1$. In this case, communication is efficient, as both firms gain from the exchange of messages: it is true that competition increases in each of the three intervals of the market size, but firms make more precise production decisions and this allows them to better extrapolate consumers' surplus.

Corollary 2.1. *When agents have complementary thresholds (i.e.* $\alpha_1 + \alpha_2 \approx 1$*), explicit collusion is more profitable than tacit collusion, even though the former is more difficult to sustain than the latter.*

Proof. See Appendix B. □

When agents are equally (or almost equally) informed and with complementary thresholds, explicit collusion is more profitable than tacit because, through communication, agents can increase their ability to understand the market demand, and hence they can set their production in order to better fit it, and to better extrapolate consumers' surplus. In other words, being able to perceive three market segments, instead of two, increases their expected profits. Our result is consistent with the experiments conducted by [Fonseca and Normann, 2012] and [Waichman et al., 2014]: they find clear evidence that communication in laboratory markets generally helps participants reach higher profits.

2.6 Welfare analysis and policy implications

In a welfare perspective, consumers are always worse off when firms collude. In fact, firms can behave in such a way to approximate the performance of a monopolist: lack of competition harms consumers (who face a decrease in their surplus), increasing the deadweight loss of society as a whole. This (intuitive) result is reflected by the fact that, when firms collude, the good is produced in smaller quantities and sold at higher prices.

[9]Literature is recently focusing on the use of algorithms as source of tacit collusion: a work by [Calvano et al., 2019] has analyzed algorithms that easily learn to play tacit collusive strategies, even without being designed to do so.

Total welfare decreases as well, meaning that the marginal decrease of consumers' surplus is greater than the marginal increase of the producers'.

In a welfare perspective, eliminating collusion would be a very important achievement; however, this is a hardly attainable. What CAs usually work on is trying to make collusion hard to reach or to sustain: they do so using not only legal and monetary punishments, but also leniency programs that incentivize firms to report existing cartels, in exchange for little to no punishment for the whistleblower.

Under this perspective, our model offers a possible solution that could prevent explicit collusion, in case agents where patient enough to attain it. Until now, we have considered the agents' thresholds as exogenous and fixed. Let us endow CA with the possibility to move α_1 and α_2, as if it could manipulate the way firms perceive the demand. This is a strong hypotheses, but it shows some curious findings.

To prevent explicit collusion, CA could move the thresholds and equalize them, making firms equally informed. In fact, when firms are equally informed, i.e. $\alpha_1 = \alpha_2$, communication becomes meaningless: with equal thresholds, it does not add any new information about the market demand, so agents will simply not communicate. Hence, in this setting there would be no formal difference between tacit and explicit collusion, as communication between firms would be useless (having the same threshold, they observe equal private signals). Moreover, being equally informed, agents have the same expected profits, and thus there are no side payments.

In the left part of Figure 2.6 we have represented the consumers' surplus when $\alpha_1 = \alpha_2$. The dashed line represents the consumers' surplus in spot competition, while the solid line represents the consumers' surplus in cartel collusion. In the right part, we plot *ex ante* profits of firm i, with $i = 1, 2$, in spot competition (dashed line) and collusion (solid) when $\alpha_1 = \alpha_2$. We plot both figures assuming, without loss of generality, that $c = 0.5$.

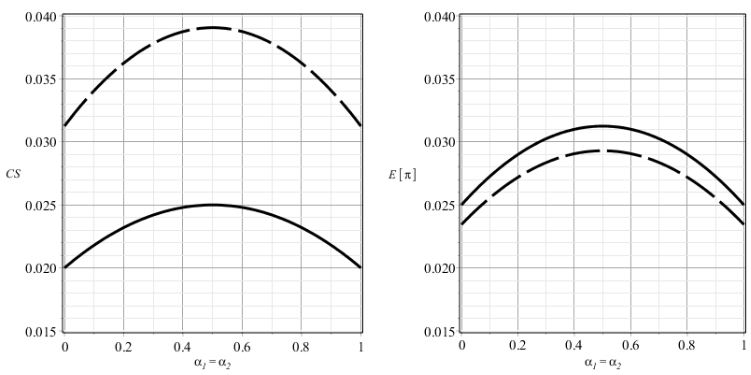


FIGURE 2.6: Consumers' surplus (left) and expected profits (right) when $\alpha_1 = \alpha_2$ (and $c = 0.5$).

With no explicit collusion possible, the only threat to competition is represented by tacit collusion: as it is shown in the left part of Figure 2.6, consumers' surplus decreases significantly if agents are able to reach an agreement. This scenario is not implausible: Figure 2.7 represents the minimum discount factor needed to attain collusion when $\alpha_1 = \alpha_2$ (again, if $c = 0.5$).

The decision where to place α_1 and α_2 raises a curious trade-off. CA could face the decision whether to (i) set $\alpha_1 = \alpha_2 = 1$: in this case, we will have two *uninformed* firms, with the coarsest possible perception of the market. This would harm the consumers, as

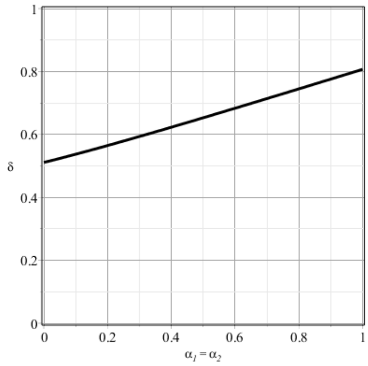


FIGURE 2.7: Minimum δ needed for collusion when $\alpha_1 = \alpha_2$ (and $c = 0.5$).

production quantities would not fine-tune the demand (see Figure 2.6).[10] Moreover, it would harm firms as well, reducing their expected profits. However, collusion would be relatively hard to attain: in fact, the minimum δ needed is equal to 0.81. Alternatively, CA could (ii) set $\alpha_1 = \alpha_2 = 0.5$, maximizing not only the consumers' surplus but also the expected profits of the firms, hence enhancing the whole total welfare but at the price of lowering the minimum discount factor necessary for collusion ($\delta = 0.66$), and thus leaving room for the possibility to collude even for relatively impatient firms.

2.7 Conclusions

In this paper, we have discussed the role of market perception on the incentives to form a cartel. Analyzing the collusion incentives of two differently informed agents playing a Cournot game, we study two types of collusion: *tacit collusion*, that we defined as the mutual decision to lower the production quantities in order to increase profits, without any exchange of relevant information between the firms, and *explicit collusion*, based on a reciprocal exchange of messages that regard the market demand. As we have defined them, only tacit collusion can be considered legal, whereas explicit collusion is prohibited because it represents a concerted practice that has as its object or effect "the prevention, restriction or distortion of competition within the internal market" which may "directly or indirectly fix [...] selling prices or [...] limit or control production" (see Article 101, TFUE). In other words, explicit collusion is illegal because it represents a secret agreement between firms aimed at eliminating or substantially reducing the uncertainty that is related to their respective commercial practices (especially in our simplified model, where the market size directly affects productions and prices). For instance, ICA has recently punished a concertation among leading captive banks and related automotive groups operating in Italy who exchanged sensitive information aimed at altering the competitive dynamics in the market of car sales. Interestingly, the investigation started thanks to the admission of a leniency applicant.[11]

We proved that, under certain conditions, explicit collusion is sustainable only through the introduction of side payments made by the *less informed* agent: a mechanism that incentivizes the *more informed* agent to share her private signal.

[10] Productions are fixed along all the market size because the signal $S_i^a = S_i^H$ has null probability to arrive for each $i = 1, 2$.

[11] See ICA Proceeding *I811 - Car sales through financing.*

The most important point of this work is that, *ceteris paribus*, explicit collusion is more difficult to attain than tacit collusion: in fact, when agents exchange their private information regarding the market demand, there are incentives to lie that rise the minimum discount factor needed to sustain this kind of collusion.

As a corollary result, we introduced a trade-off that regards producers with complementary categories: explicit collusion increases profits over the level given by tacit collusion, however the former is both more difficult to attain and more detectable than the latter.

Finally, in a welfare perspective, a Competition Authority that seeks to maximize consumers' surplus would find optimal to have two *best-informed* firms, i.e. two firms that divide the low- and the high-market segments in equal length. This comes with the cost that tacit collusion would still be possible, and not difficult to attain.

Our model neglects the role of CAs, i.e. we have assumed no costs for entering (or maintaining) a cartel agreement which could be object of future extensions. As explicit collusion is unlawful, we might study the process by which firms are deterred. Hence, we could add to the model a probability of an investigation and, conditional on an investigation, a probability of conviction. We believe these features will not significantly change our results, and possibly just increase their consistency.

Another possible expansion of the model would be endogenizing categorization and thresholds, in order to study the agents' categorical thinking and their development in a dynamic setting. It would be interesting to assess whether there might exist equilibrium categorizations that may emerge in the long-run, and how they relate with the attainability of collusion.

Moreover, it would be also interesting to test our hypothesis through an experiment, and see whether our findings could be supported by real-life examples, increasing the potential empirical relevance of this work.

Other assumptions could be relaxed: number of categories, equal discount factors and cost functions, Nash reversion as punishment. We believe this would not contribute much to the analysis, adding a big cost in terms of tractability.

Appendix A

Just for this section, as formulas are more convenient to be reported, we substituted $\alpha_1 = a$ and $\alpha_2 = b$.

$$x_1^{L,spot} = 1/2\, \frac{4\,a^2bc^2 + 10\,a^2bc - 2\,ab^2c - 4\,abc^2 + 6\,a^2b - 3\,ab^2 - 10\,abc + 2\,b^2c - 6\,ab + 3\,b^2 + a - b}{8\,abc^3 + 28\,abc^2 - 8\,bc^3 + 30\,abc - 28\,bc^2 + 9\,ab + 2\,ac - 32\,bc + 3\,a - 12\,b}$$

$$x_1^{H,spot} = 1/2\, \frac{4\,a^2bc^2 + 10\,a^2bc - 2\,ab^2c + 6\,a^2b - 3\,ab^2 - 4\,bc^2 - 8\,bc + a - 4\,b}{8\,abc^3 + 28\,abc^2 - 8\,bc^3 + 30\,abc - 28\,bc^2 + 9\,ab + 2\,ac - 32\,bc + 3\,a - 12\,b}$$

$$x_2^{L,spot} = -1/2\, \frac{-4\,ab^2c^2 + 2\,a^2bc - 10\,ab^2c + 4\,b^2c^2 + 3\,a^2b - 2\,a^2c - 6\,ab^2}{8\,abc^3 + 28\,abc^2 - 8\,bc^3 + 30\,abc - 28\,bc^2 + 9\,ab + 2\,ac - 32\,bc + 3\,a - 12\,b} - $$
$$- 1/2\, \frac{+10\,b^2c - 3\,a^2 + 2\,ac + 6\,b^2 - 2\,bc + 2\,a - 2\,b}{8\,abc^3 + 28\,abc^2 - 8\,bc^3 + 30\,abc - 28\,bc^2 + 9\,ab + 2\,ac - 32\,bc + 3\,a - 12\,b}$$

$$x_2^{H,spot} = -1/2\, \frac{-4\,ab^2c^2 + 2\,a^2bc - 10\,ab^2c - 4\,abc^2 + 4\,b^2c^2 + 3\,a^2b - 6\,ab^2 - 10\,abc + 10\,b^2c}{8\,abc^3 + 28\,abc^2 - 8\,bc^3 + 30\,abc - 28\,bc^2 + 9\,ab + 2\,ac - 32\,bc + 3\,a - 12\,b} - $$
$$- 1/2\, \frac{+4\,bc^2 - 6\,ab + 6\,b^2 + 8\,bc - a + 4\,b}{8\,abc^3 + 28\,abc^2 - 8\,bc^3 + 30\,abc - 28\,bc^2 + 9\,ab + 2\,ac - 32\,bc + 3\,a - 12\,b}$$

$$x_1^{L,tacit} = 1/4\, \frac{a^2bc^2 + 3\,a^2bc - ab^2c - abc^2 + 2\,a^2b - 2\,ab^2 - 3\,abc + b^2c - 2\,ab + 2\,b^2 + a - b}{abc^3 + 4\,abc^2 - bc^3 + 4\,abc - 4\,bc^2 + ac - 5\,bc + 2\,a - 2\,b}$$

$$x_1^{H,tacit} = 1/4\, \frac{a^2bc^2 + 3\,a^2bc - ab^2c + 2\,a^2b - 2\,ab^2 - bc^2 - 2\,bc + a - b}{abc^3 + 4\,abc^2 - bc^3 + 4\,abc - 4\,bc^2 + ac - 5\,bc + 2\,a - 2\,b}$$

$$x_2^{L,tacit} = -1/4\, \frac{-ab^2c^2 + a^2bc - 3\,ab^2c + b^2c^2 + 2\,a^2b - a^2c - 2\,ab^2 + 3\,b^2c - 2\,a^2 + ac + 2\,b^2 - bc + a - b}{abc^3 + 4\,abc^2 - bc^3 + 4\,abc - 4\,bc^2 + ac - 5\,bc + 2\,a - 2\,b}$$

$$x_2^{H,tacit} = -1/4\, \frac{-ab^2c^2 + a^2bc - 3\,ab^2c - abc^2 + b^2c^2 + 2\,a^2b - 2\,ab^2 - 3\,abc}{abc^3 + 4\,abc^2 - bc^3 + 4\,abc - 4\,bc^2 + ac - 5\,bc + 2\,a - 2\,b} - $$
$$- 1/4\, \frac{+3\,b^2c + bc^2 - 2\,ab + 2\,b^2 + 2\,bc - a + b}{abc^3 + 4\,abc^2 - bc^3 + 4\,abc - 4\,bc^2 + ac - 5\,bc + 2\,a - 2\,b}$$

$$x_1^{L,expl} = x_2^{L,expl} = 1/4\, \frac{a}{b+2}$$

$$x_1^{M,expl} = x_2^{M,expl} = 1/4\, \frac{a+b}{c+2}$$

$$x_1^{H,expl} = x_2^{H,expl} = 1/4\, \frac{a+1}{b+2}$$

Appendix B

As it is possible to grasp from Appendix A, the functions we are dealing with (and especially the discount factors) are complex and difficult to simplify. Moreover, we cannot find clear patterns (e.g. discount factors are not always increasing in c, or in the difference $\alpha_1 - \alpha_2$). Hence, we need to use graphical analysis in order to prove the four proposition. However, as discount factors and profit function depend on three parameters (α_1, α_2, and c), we are forced to fix one of these parameters (for instance, c) and plot the discount factors as function of the other two. In some cases we cannot use a 3D plot because the functions present asymptotes outside their domain (for instance, the function $\delta_{2,W}^{H,exp}$ cannot be plotted in $\alpha_1, \alpha_2 \in [0,1]^2$ and for a fixed value of c because because it has asymptotes outside $r = W$). In this cases, we adopt two different strategies depending on the area r that we are considering. When we are considering either $r = W$ or $r = N$, we fix not only c, but also another parameter (either α_1 or α_2, depending on the area we are considering), and plot the discount factor as function of the remaining parameter α_i, which varies along the limits of the considered region r: let $\underline{\alpha}_i$ and $\bar{\alpha}_i$ denote these limits, that change according to the value of α_j and c. For instance, in $r = W$, when $\alpha_1 = 0.14$ and $c = 0.3$, we have that $\underline{\alpha}_2 = 0.346$ and $\bar{\alpha}_2 = 0.636$; or, when $r = N$, when $\alpha_2 = 0.86$ and $c = 0.3$ we have that $\underline{\alpha}_1 = 0.364$ and $\bar{\alpha}_1 = 654$. When we are considering $r = C$, we look at the two diagonals of the square $\alpha_1 \times \alpha_2 \in [0,1]^2$ and slight switches, such as $\alpha_2 = \alpha_1 + 0.1$ or $\alpha_1 + \alpha_2 = 0.95$, as the central area mainly unfolds around them.

For every parameter that we fix to an arbitrary value (for instance $\alpha_1 = 0.07$), we could have chosen a different one (within the limits of the considered area r) and results would have not changed. We restrict our choice to only two representative values, because considering more would have excessively prolonged the Appendix. In Table 2.1 we summarize the propositions and for which values of α_1 and α_2 we proved them. Note that every statement is proved for different values of c: we use $c = \{0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 0.99\}$ (hereafter, we shorten this notation to $c = \{0.1, 0.2, 0.3, ..., 0.99\}$).

TABLE 2.1: Summary of proofs.

Proposition	**Statement**	$\boldsymbol{\alpha_1, \alpha_2}$
2.1	$\delta_1^{H,tacit} = \max\{\delta_1^{H,tacit}, \delta_1^{L,tacit}\}$	$\forall\ \alpha_1, \alpha_2 \in [0,1]^2$
	$\delta_2^{H,tacit} = \max\{\delta_2^{H,tacit}, \delta_2^{L,tacit}\}$	$\forall\ \alpha_1, \alpha_2 \in [0,1]^2$
2.2	$\delta_{1,W}^{H,exp} = \max\{\delta_{1,W}^{H,exp}, \delta_{1,W}^{M,exp}, \delta_{1,W}^{L,exp}\}$	$\forall\ \alpha_1, \alpha_2 \in [0,1]^2$
	$\delta_{2,W}^{H,exp} = \max\{\delta_{2,W}^{H,exp}, \delta_{2,W}^{M,exp}, \delta_{2,W}^{L,exp}\}$	$\alpha_1 = 0.07, \alpha_2 \in (\underline{\alpha}_2, \bar{\alpha}_2)$
		$\alpha_1 = 0.14, \alpha_2 \in (\underline{\alpha}_2, \bar{\alpha}_2)$
	$\delta_{1,N}^{H,exp} = \max\{\delta_{1,N}^{H,exp}, \delta_{1,N}^{M,exp}, \delta_{1,N}^{L,exp}\}$	$\alpha_2 = 0.93, \alpha_1 \in (\underline{\alpha}_1, \bar{\alpha}_1)$
		$\alpha_2 = 0.86, \alpha_1 \in (\underline{\alpha}_1, \bar{\alpha}_1)$
	$\delta_{2,N}^{H,exp} = \max\{\delta_{2,N}^{H,exp}, \delta_{2,N}^{M,exp}, \delta_{2,N}^{L,exp}\}$	$\forall\ \alpha_1, \alpha_2 \in [0,1]^2$
	$\delta_{1,C}^{H,exp} = \max\{\delta_{1,C}^{H,exp}, \delta_{1,C}^{M,exp}, \delta_{1,C}^{L,exp}\}$	$\{\alpha_2 = \alpha_1, \alpha_2 = \alpha_1 + 0.1\}$
		$\alpha_1 + \alpha_2 = \{0.95, 1.05\}$
	$\delta_{2,C}^{H,exp} = \max\{\delta_{2,C}^{H,exp}, \delta_{2,C}^{M,exp}, \delta_{2,C}^{L,exp}\}$	$\{\alpha_2 = \alpha_1, \alpha_2 = \alpha_1 + 0.1\}$
		$\alpha_1 + \alpha_2 = \{0.95, 1.05\}$
2.3	$\max\{\delta_{1,W}^{H,expl}, \delta_{2,W}^{H,expl}\} \geqslant \max\{\delta_1^{H,tacit}, \delta_2^{H,tacit}\}$	$\alpha_1 = 0.07, \alpha_2 \in (\underline{\alpha}_2, \bar{\alpha}_2))$
		$\alpha_1 = 0.14, \alpha_2 \in (\underline{\alpha}_2, \bar{\alpha}_2)$
	$\max\{\delta_{1,N}^{H,expl}, \delta_{2,N}^{H,expl}\} \geqslant \max\{\delta_1^{H,tacit}, \delta_2^{H,tacit}\}$	$\alpha_2 = 0.93, \alpha_1 \in (\underline{\alpha}_1, \bar{\alpha}_1)$
		$\alpha_2 = 0.86, \alpha_1 \in (\underline{\alpha}_1, \bar{\alpha}_1)$
	$\max\{\delta_{1,C}^{H,expl}, \delta_{2,C}^{H,expl}\} \geqslant \max\{\delta_1^{H,tacit}, \delta_2^{H,tacit}\}$	$\alpha_1 + \alpha_2 = 0.95$
		$\alpha_1 + \alpha_2 = 1.05$

Proposition 2.1. *In tacit collusion, $\delta_i^{H,tacit} = \max\left\{\delta_i^{H,tacit}, \delta_i^{L,tacit}\right\}$ represents the highest minimum discount factor, and hence the only one which is binding, for each agent $i = 1, 2$.*

Proof. We show that $\delta_i^{H,tacit} = \max\left\{\delta_i^{H,tacit}, \delta_i^{L,tacit}\right\}$ holds for each $i = 1, 2$, for $\alpha_1, \alpha_2 \in [0,1]^2$ (with $\alpha_1 \leqslant \alpha_2$), and for $c \in (0,1)$. We first prove $\delta_1^{H,tacit} \geqslant \delta_1^{L,tacit}$ in Figure 2.8. We plot $\delta_1^{H,tacit}$ in red and $\delta_1^{L,tacit}$ in blue as function of α_1 and α_2 for different values of c. In particular, from up-left to down-right, we plot the two functions for ten different values of c, i.e. $c = \{0.1, 0.2, 0.3, ..., 0.99\}$.

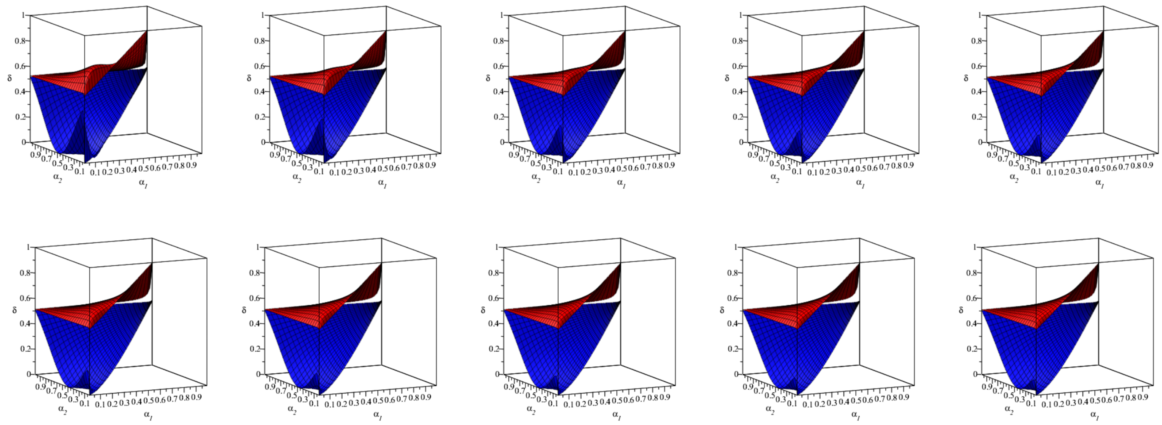


FIGURE 2.8: $\delta_1^{H,tacit}$ (blue) is greater than $\delta_1^{L,tacit}$ (red), for different values of c.

For instance, when $\alpha_1 = 0.3$, $\alpha_2 = 0.4$, and $c = 0.3$, $\delta_1^{H,tacit} = 0.649$ and $\delta_1^{L,tacit} = 0.075$.

We can apply the same technique to prove $\delta_2^{H,tacit} \geqslant \delta_2^{L,tacit}$. In Figure 2.9 we plot $\delta_2^{H,tacit}$ in red and $\delta_2^{L,tacit}$ in blue as function of α_1 and α_2 for $c = \{0.1, 0.2, 0.3, ..., .99\}$.

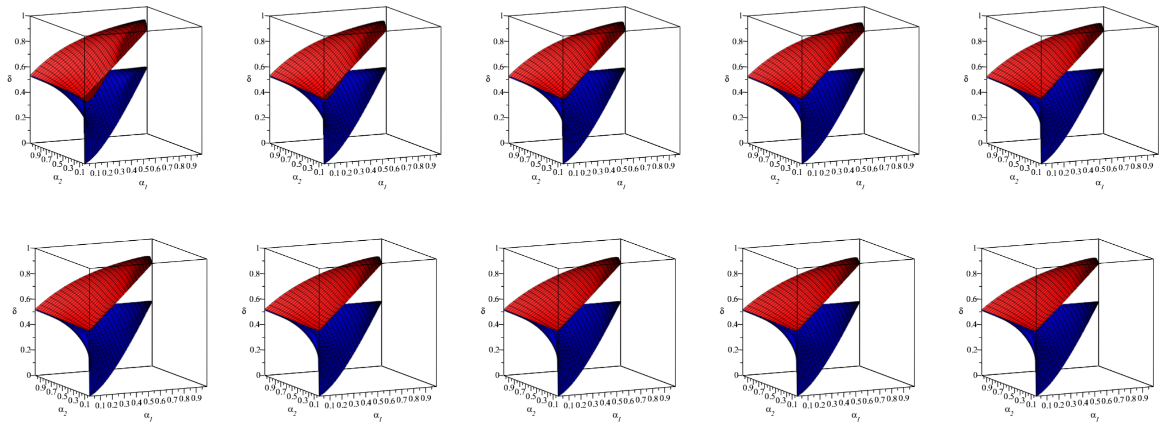
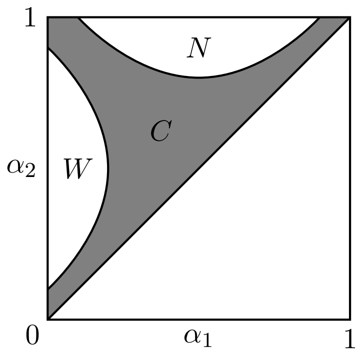


FIGURE 2.9: $\delta_2^{H,tacit}$ (blue) is greater than $\delta_2^{L,tacit}$ (red), for different values of c.

For instance, when $\alpha_1 = 0.3$, $\alpha_2 = 0.4$, and $c = 0.3$, $\delta_2^{H,tacit} = 0.572$ and $\delta_2^{L,tacit} = 0.202$. □

Proposition 2.2. *In explicit collusion,* $\delta_{i,r}^{H,expl} = \max\left\{\delta_{i,r}^{H,expl}, \delta_{i,r}^{M,expl}, \delta_{i,r}^{L,expl}\right\}$ *represents the highest minimum discount factor for each agent* $i = 1, 2$ *and each region* $r = W, N, C$.

Proof. To ease the readability of the proof, we report here Figure 2.4.



We first focus on the west area of Figure 2.4, i.e. $r = W$. In Figure 2.10 we first prove that $\delta_{1,W}^{H,exp} = \max\left\{\delta_{1,W}^{H,exp}, \delta_{1,W}^{M,exp}, \delta_{1,W}^{L,exp}\right\}$. We plot $\delta_{1,W}^{H,exp}$ in red, $\delta_{1,W}^{M,exp}$ in blue, and $\delta_{1,W}^{L,exp}$ in green, as function of α_1 and α_2 for different values of c. In particular, from up-left to down-right, we plot the three functions for $c = \{0.1, 0.2, 0.3, ..., 0.99\}$.

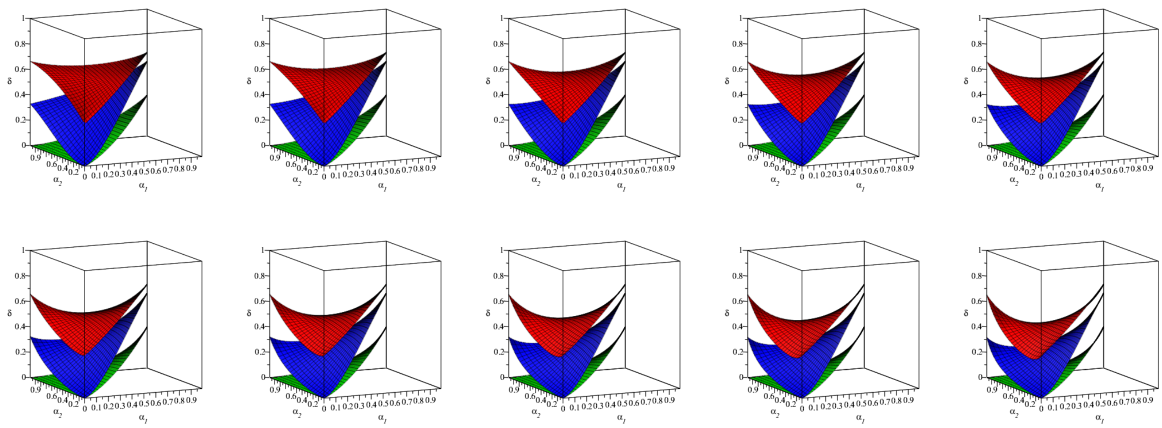


FIGURE 2.10: $\delta_{1,W}^{H,exp}$ (red) is greater than $\delta_{1,W}^{M,exp}$ (blue) and $\delta_{1,W}^{L,exp}$ (green), for different values of c.

For instance, when $\alpha_1 = 0.14$, $\alpha_2 = 0.4$, and $c = 0.3$, $\delta_{1,W}^{H,exp} = 0.455$ and $\delta_{1,W}^{M,exp} = 0.104$, and $\delta_{1,W}^{L,exp} = 0.008$.

To prove $\delta_{2,W}^{H,exp} = \max\left\{\delta_{2,W}^{H,exp}, \delta_{2,W}^{M,exp}, \delta_{2,W}^{L,exp}\right\}$, the nature of the functions does not allow us to use 3D plots. Hence, we fix α_1. In Figure 2.11, we plot the three discount factors $\delta_{2,W}^{H,exp}$ in red, $\delta_{2,W}^{M,exp}$ in blue, and $\delta_{2,W}^{L,exp}$ in green, as function of α_2 for ten different values of c, i.e. $c = \{0.1, 0.2, 0.3, ..., 0.99\}$. In the first ten plots (from up-left to down-right) we fix $\alpha_1 = 0.07$, while in the other ten we use $\alpha_1 = 0.14$.

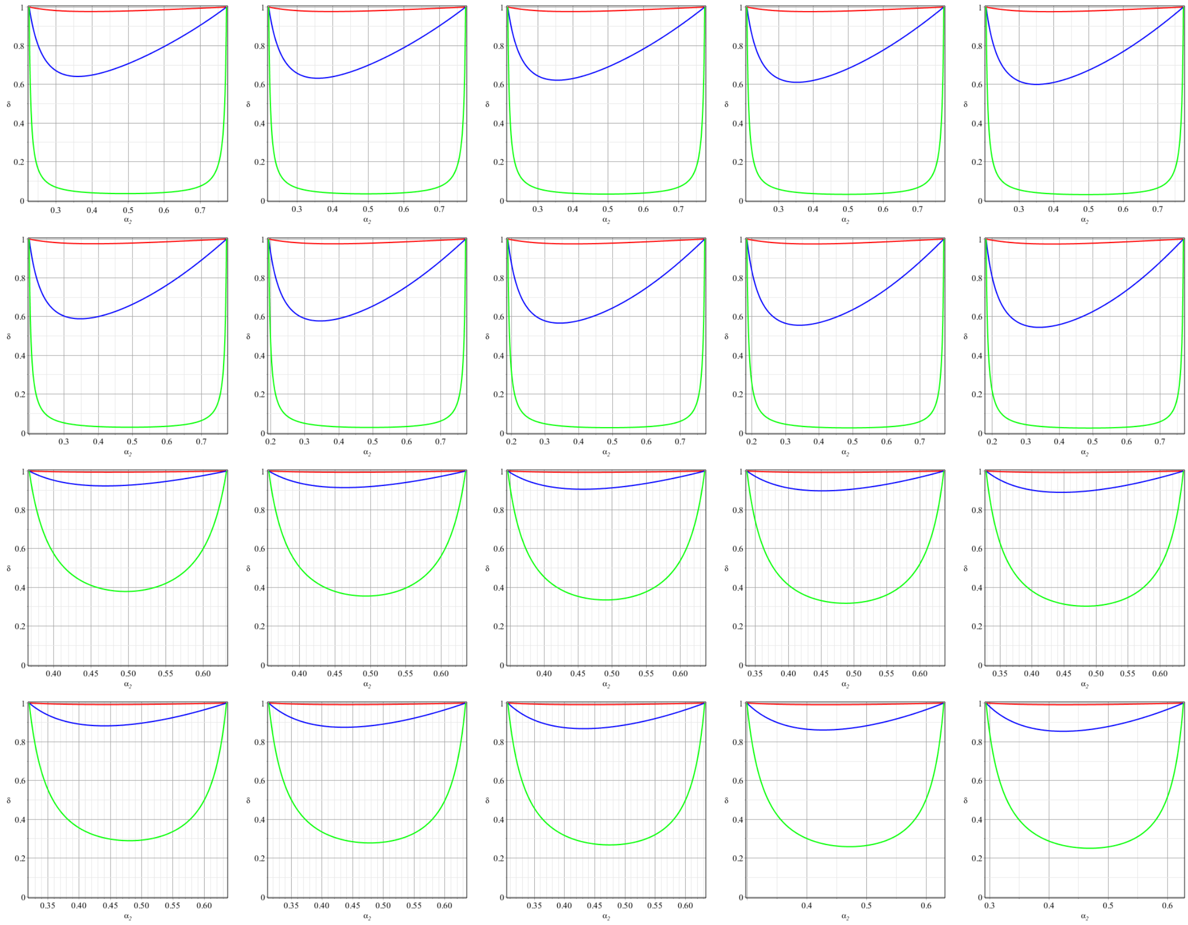


FIGURE 2.11: $\delta_{2,W}^{H,exp}$ (red) is greater than $\delta_{2,W}^{M,exp}$ (blue) and $\delta_{2,W}^{L,exp}$ (green), for different values of c.

For instance, when $\alpha_1 = 0.14$, $\alpha_2 = 0.4$, and $c = 0.3$, we have that $\delta_{2,W}^{H,exp} = 0.995$, $\delta_{2,W}^{M,exp} = 0.925$, and $\delta_{2,W}^{L,exp} = 0.453$.

Now we focus on the north area $r = N$. To prove $\delta_{1,N}^{H,exp} = \max\left\{\delta_{1,N}^{H,exp}, \delta_{1,N}^{M,exp}, \delta_{1,N}^{L,exp}\right\}$, this time we fix α_2. In Figure 2.12, we plot the three discount factors $\delta_{1,N}^{H,exp}$ in red, $\delta_{1,N}^{M,exp}$ in blue, and $\delta_{1,N}^{L,exp}$ in green, as function of α_1 for $c = \{0.1, 0.2, 0.3, ..., 0.99\}$ from up-left to down-right. In the first ten plots we fix $\alpha_2 = 0.93$, while in the other ten we use $\alpha_2 = 0.86$.

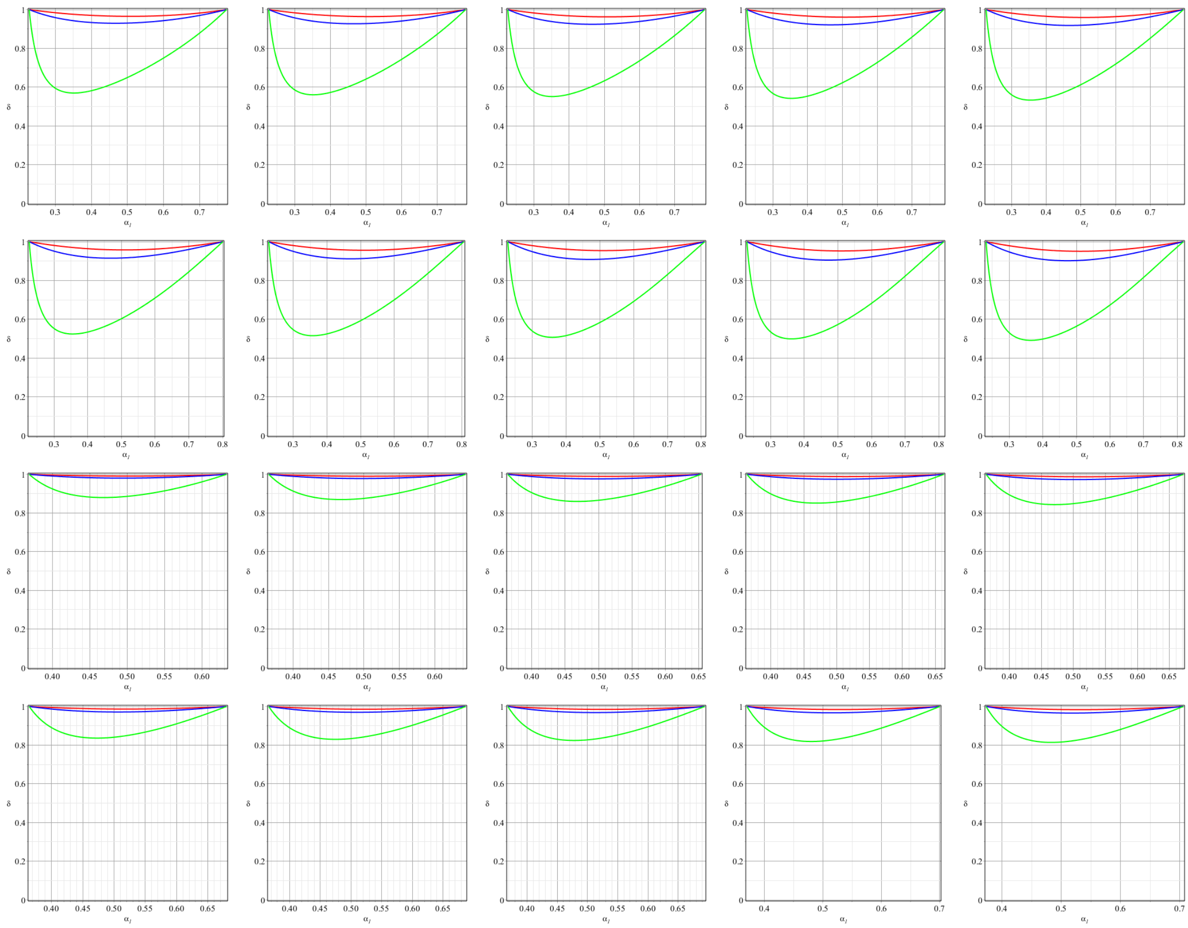


FIGURE 2.12: $\delta_{1,N}^{H,exp}$ (red) is greater than $\delta_{1,N}^{M,exp}$ (blue) and $\delta_{1,N}^{L,exp}$ (green), for different values of c.

For instance, when $\alpha_1 = 0.4$, $\alpha_2 = 0.86$, $c = 0.3$, we have that $\delta_{1,N}^{H,exp} = 0.995$, $\delta_{1,N}^{M,exp} = 0.988$, and $\delta_{1,N}^{L,exp} = 0.905$

In Figure 2.13 we prove that $\delta_{2,N}^{H,exp} = \max\left\{\delta_{2,N}^{H,exp}, \delta_{2,N}^{M,exp}, \delta_{2,N}^{L,exp}\right\}$. We plot $\delta_{2,N}^{H,exp}$ in red, $\delta_{2,N}^{M,exp}$ in blue, and $\delta_{2,N}^{L,exp}$ in green as function of α_1 and α_2 for different values of c. In particular, we plot the three functions for $c = \{0.1, 0.2, 0.3, ..., 0.99\}$ (from up-left to down-right).

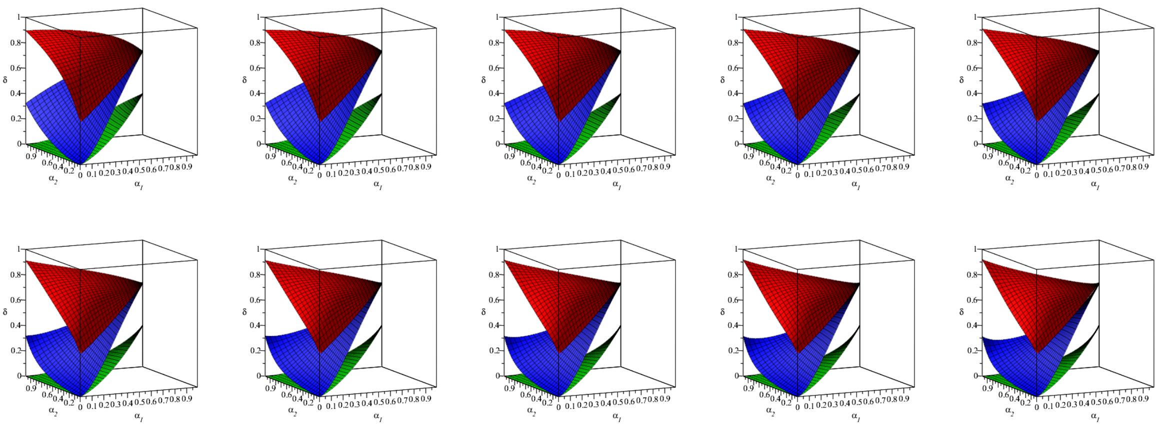


FIGURE 2.13: $\delta_{2,N}^{H,exp}$ (red) is greater than $\delta_{2,N}^{M,exp}$ (blue) and $\delta_{2,N}^{L,exp}$ (green), for different values of c.

For instance, when $\alpha_1 = 0.4$, $\alpha_2 = 0.86$, and $c = 0.3$, $\delta_{2,N}^{H,exp} = 0.806$ and $\delta_{2,N}^{M,exp} = 0.401$ and $\delta_{2,N}^{L,exp} = 0.063$.

Now we focus on the central grey area ($r = C$) that mainly unfolds around the two diagonals. For this area, as representative cases, we only consider the main diagonal $\alpha_2 = \alpha_1$ (and one slight switch $\alpha_2 = \alpha_1 + \varepsilon$) and two slight switches around the antidiagonal $\alpha_1 + \alpha_2 = 1$ (i.e. we consider $\alpha_1 + \alpha_2 = 1 \pm \frac{\varepsilon}{2}$) of the square space $\alpha_1, \alpha_2 = [0, 1]^2$.

In Figure 2.14 we first prove that $\delta_{1,C}^{H,exp} = \max\left\{\delta_{1,C}^{H,exp}, \delta_{1,C}^{M,exp}, \delta_{1,C}^{L,exp}\right\}$. We plot $\delta_{1,C}^{H,exp}$ in red, $\delta_{1,C}^{M,exp}$ in blue, and $\delta_{1,C}^{L,exp}$ in green as function of α for $c = \{0.1, 0.2, 0.3, ..., 0.99\}$. In the first ten squares, we represent $\delta_{1,C}^{H,exp}$, $\delta_{1,C}^{M,exp}$, and $\delta_{1,C}^{L,exp}$ for $\alpha_2 = \alpha_1$ (solid lines, in order to sketch the diagonal), and then for $\alpha_2 = \alpha_1 + \varepsilon$ (dashed lines, in order to sketch a slight switch from the diagonal towards the north-west corner of the square $\alpha_1, \alpha_2 = [0, 1]^2$), where we arbitrarily set $\varepsilon = 0.1$ (even if negligible, there is a slight change as c increases). In the other ten squares, we represent $\delta_{1,C}^{H,exp}$, $\delta_{1,C}^{M,exp}$, and $\delta_{1,C}^{L,exp}$ for $\alpha_1 + \alpha_2 = 0.95$ (solid lines) and for $\alpha_1 + \alpha_2 = 1.05$ (dashed lines) as to sketch two slight switches from the antidiagonal.

Hence, when analyzing $\delta_{1,C}^{H,exp}$, $\delta_{1,C}^{M,exp}$, and $\delta_{1,C}^{L,exp}$ we have to keep in mind that we have to compare solid lines with solid lines and dashed lines with dashed lines.

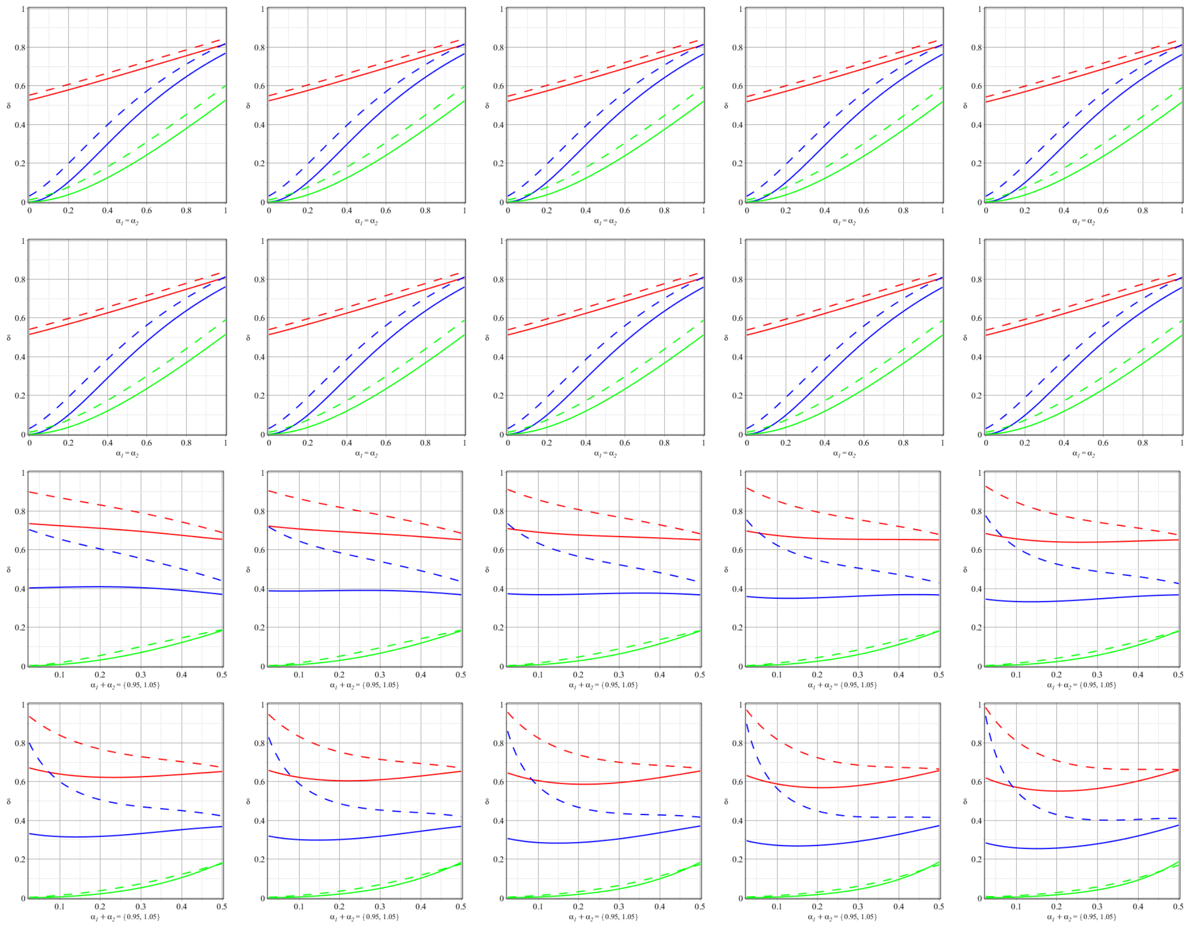


FIGURE 2.14: $\delta_{1,C}^{H,exp}$ (red) is greater than $\delta_{1,C}^{M,exp}$ (blue) and $\delta_{1,C}^{L,exp}$ (green), for different values of c.

For instance, when $\alpha_1 = 0.3$, $\alpha_2 = 0.4$, and $c = 0.3$, $\delta_{1,C}^{H,exp} = 0.567$ and $\delta_{1,C}^{M,exp} = 0.211$, and $\delta_{1,C}^{L,exp} = 0.057$.

Similarly, in Figure 2.15 we prove that $\delta_{2,C}^{H,exp} = \max\left\{\delta_{2,C}^{H,exp}, \delta_{2,C}^{M,exp}, \delta_{2,C}^{L,exp}\right\}$. We plot $\delta_{2,C}^{H,exp}$ in red, $\delta_{2,C}^{M,exp}$ in blue, and $\delta_{2,C}^{L,exp}$ in green as function of α for $c = \{0.1, 0.2, 0.3, ..., .99\}$. As before, in the first ten squares, we represent $\delta_{2,C}^{H,exp}$, $\delta_{2,C}^{M,exp}$, and $\delta_{2,C}^{L,exp}$ for $\alpha_2 = \alpha_1$ (solid lines, in order to sketch the diagonal), and then for $\alpha_2 = \alpha_1 + \varepsilon$ (dashed lines, in order to sketch a slight switch from the diagonal towards the north-west corner of the square), where we arbitrarily set $\varepsilon = 0.1$. In the other ten squares, we represent $\delta_{2,C}^{H,exp}$, $\delta_{2,C}^{M,exp}$, and $\delta_{2,C}^{L,exp}$ for $\alpha_1 + \alpha_2 = 0.95$ (solid lines) and for $\alpha_1 + \alpha_2 = 1.05$ (dashed lines) as to sketch two slight switches from the antidiagonal.

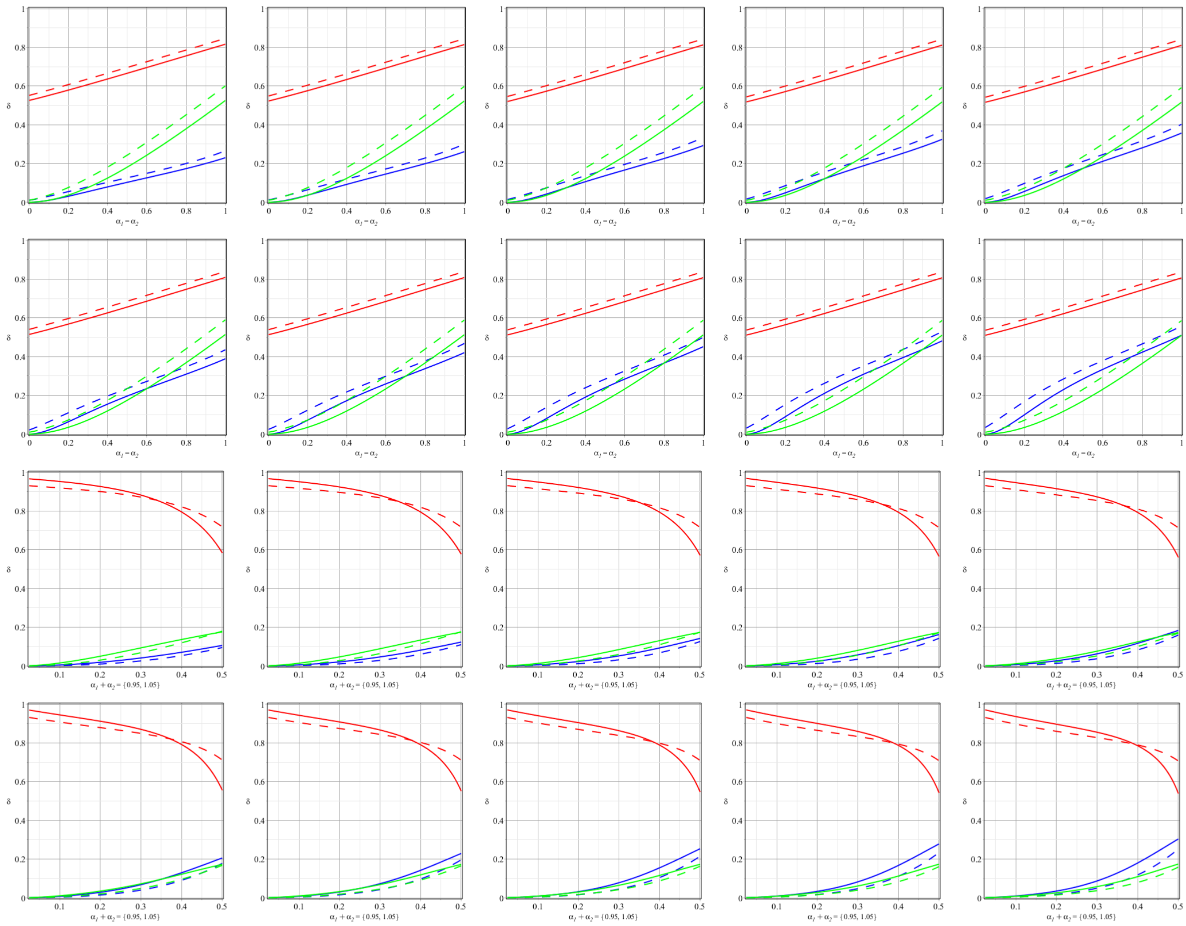


FIGURE 2.15: $\delta_{2,C}^{H,exp}$ (red) is greater than $\delta_{2,C}^{M,exp}$ (blue) and $\delta_{2,C}^{L,exp}$ (green), for different values of c.

For instance, when $\alpha_1 = 0.3$, $\alpha_2 = 0.4$, and $c = 0.3$, $\delta_{2,C}^{H,exp} = 0.789$ and $\delta_{2,C}^{M,exp} = 0.088$, and $\delta_{2,C}^{L,exp} = 0.101$.

□

Proposition 2.3. *Tacit collusion is easier to attain than explicit collusion: there are values of the common discount factor such that the former is an equilibrium while the latter is not.*

Proof. To prove this proposition we have to show that

$$\max\left\{\delta_{1,r}^{H,expl}, \delta_{2,r}^{H,expl}\right\} \geqslant \max\left\{\delta_{1}^{H,tacit}, \delta_{2}^{H,tacit}\right\}$$

for each $r = W, N, C$. First, consider the west area $r = W$ and prove that

$$\max\left\{\delta_{1,W}^{H,expl}, \delta_{2,W}^{H,expl}\right\} \geqslant \max\left\{\delta_{1}^{H,tacit}, \delta_{2}^{H,tacit}\right\}$$

With a procedure that resembles the one used in the previous proof, to plot these functions we fix α_1 and let α_2 be included within the intervals of the west area. In Figure 2.16 we plot $\delta_{1,W}^{H,expl}$ and $\delta_{2,W}^{H,expl}$ in red, and $\delta_{1}^{H,tacit}$ and $\delta_{2}^{H,tacit}$ in green (we do not distinguish between pairs) for different values of $c = \{0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 0.99\}$. We fix $\alpha_1 = 0.07$ in the first ten plots (one for each value of c), and $\alpha_1 = 0.14$ in the others and let α_2 vary along the limits of the west area.

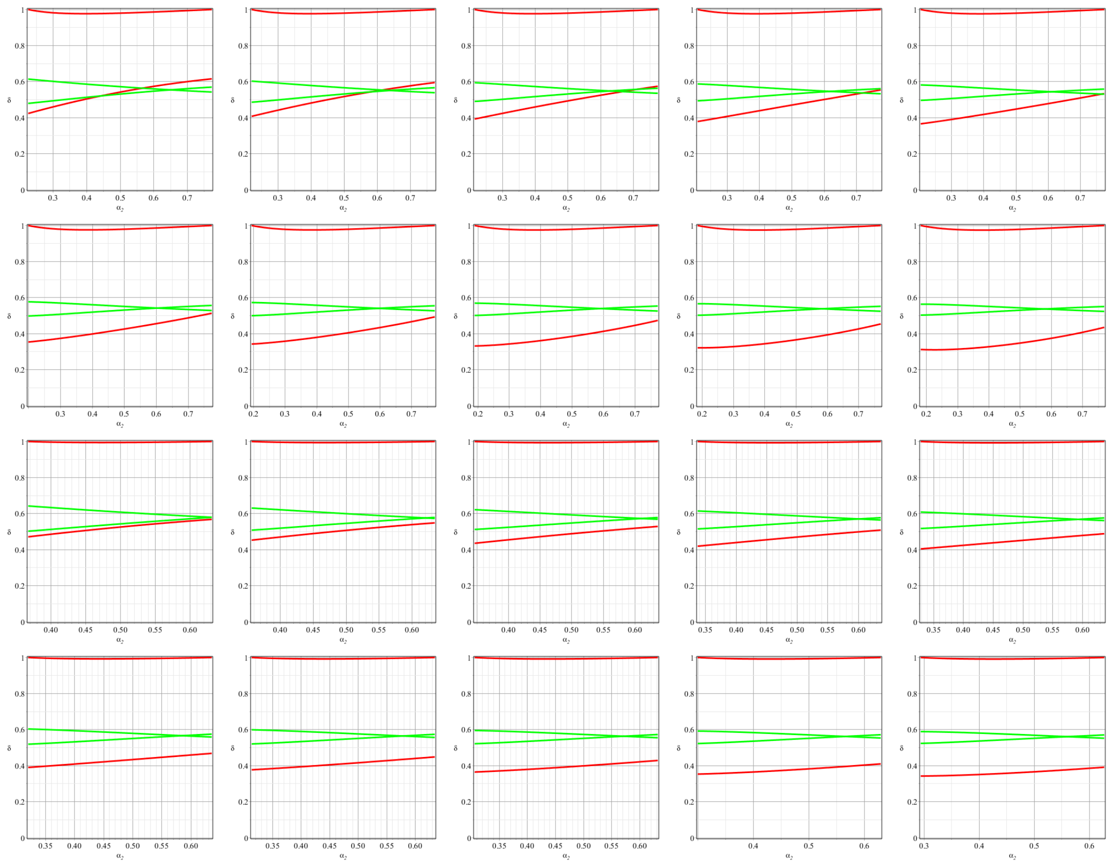


FIGURE 2.16: $\max\left\{\delta_{1,W}^{H,expl}, \delta_{2,W}^{H,expl}\right\}$ (both functions in red) is greater than $\max\left\{\delta_{1}^{H,tacit}, \delta_{2}^{H,tacit}\right\}$ (both functions in green), for different values of c.

For instance, when $\alpha_1 = 0.14$, $\alpha_2 = 0.4$, and $c = 0.3$, we have that $\max\left\{\delta_{1,W}^{H,expl}, \delta_{2,W}^{H,expl}\right\} = \max\left\{0.455, 0.995\right\} = 0.995$, while $\max\left\{\delta_{1}^{H,tacit}, \delta_{2}^{H,tacit}\right\} = \max\left\{0.612, 0.524\right\} = 0.612$.

We now consider the north area ($r = N$) and prove that

$$\max\left\{\delta_{1,N}^{H,expl}, \delta_{2,N}^{H,expl}\right\} \geqslant \max\left\{\delta_1^{H,tacit}, \delta_2^{H,tacit}\right\}$$

We plot $\delta_{1,N}^{H,expl}$ and $\delta_{2,N}^{H,expl}$ in red, while we plot $\delta_1^{H,tacit}$ and $\delta_2^{H,tacit}$ in green. In the first ten plots (one for each value of c), we fix $\alpha_2 = 0.93$ and let α_1 vary along the limits of the north area for $c = \{0.1, 0.2, 0.3, ..., 0.99\}$. In the other ten plots, we fix $\alpha_2 = 0.86$.

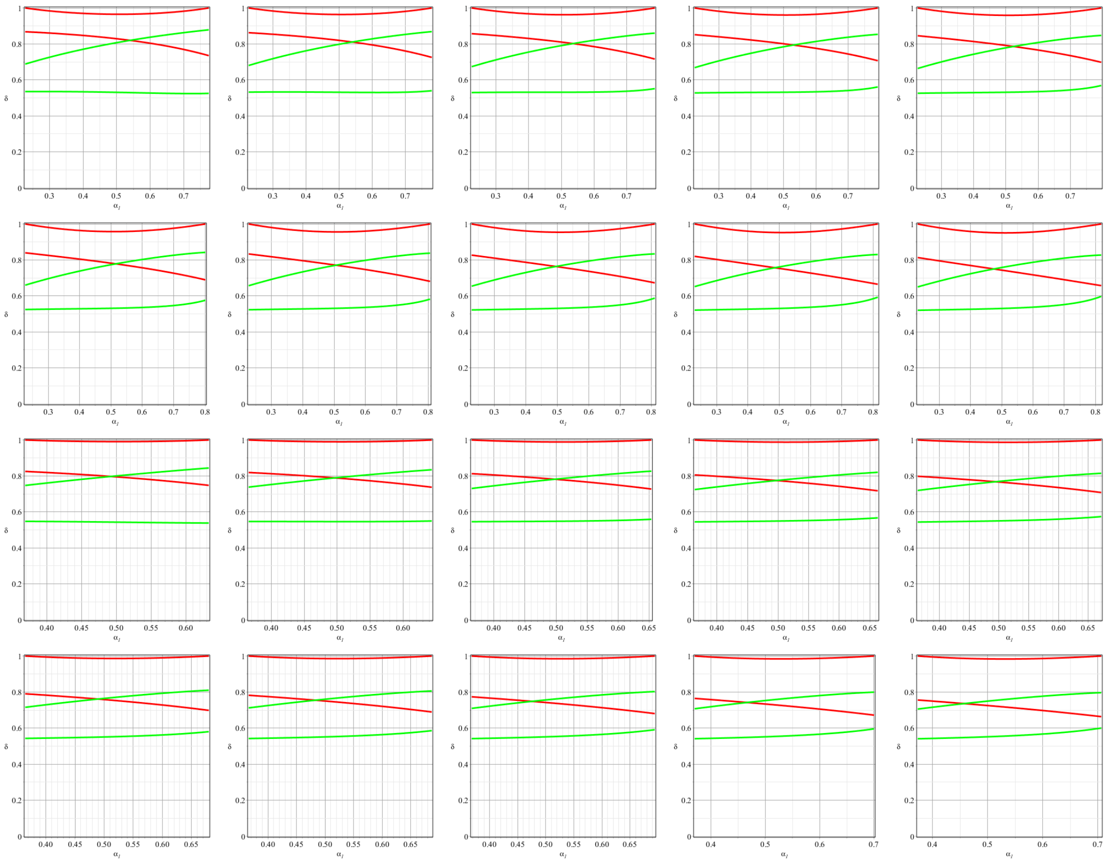


FIGURE 2.17: $\max\left\{\delta_{1,N}^{H,expl}, \delta_{2,N}^{H,expl}\right\}$ (both functions in red) is greater than $\max\left\{\delta_1^{H,tacit}, \delta_2^{H,tacit}\right\}$ (both functions in green), for different values of c.

For instance, when $\alpha_1 = 0.4$, $\alpha_2 = 0.86$, and $c = 0.3$, we have that $\max\left\{\delta_{1,N}^{H,expl}, \delta_{2,N}^{H,expl}\right\} = \max\left\{0.995, 0.806\right\} = 0.995$, while $\max\left\{\delta_1^{H,tacit}, \delta_2^{H,tacit}\right\} = \max\left\{0.546, 0.745\right\} = 0.745$.

Lastly, we consider the central area, $r = C$. In Figure 2.18 we prove that $\max\left\{\delta_{1,C}^{H,expl}, \delta_{2,C}^{H,expl}\right\} \geqslant \max\left\{\delta_1^{H,tacit}, \delta_2^{H,tacit}\right\}$. We plot $\delta_{1,C}^{H,expl}$ and $\delta_{2,C}^{H,expl}$ in red, while we plot $\delta_1^{H,tacit}$ and $\delta_2^{H,tacit}$ in green. We plot them for ten different values of c, i.e. $c = \{0.1, 0.2, 0.3, ..., 0.99\}$. In the first ten plots (one for each value of c), we let $\alpha_1 + \alpha_2 = 0.95$; in the other ten plots, we let $\alpha_1 + \alpha_2 = 1.05$.

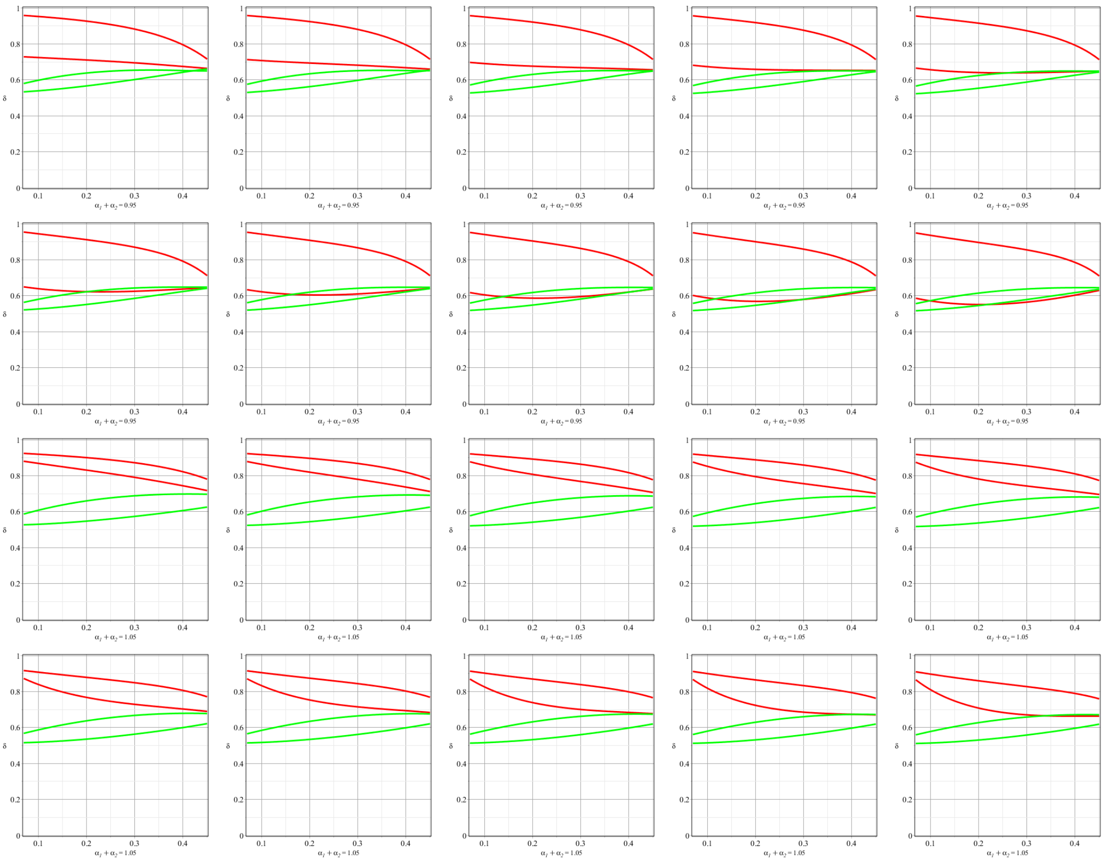


FIGURE 2.18: $\max\left\{\delta_{1,C}^{H,expl}, \delta_{2,C}^{H,expl}\right\}$ (both functions in red) is greater than $\max\left\{\delta_1^{H,tacit}, \delta_2^{H,tacit}\right\}$ (both functions in green), for different values of c.

For instance, when $\alpha_1 = 0.4$, $\alpha_2 = 0.65$, and $c = 0.3$, we have that $\max\left\{\delta_{1,C}^{H,expl}, \delta_{2,C}^{H,expl}\right\} = \max\left\{0.729, 0.817\right\} = 0.995$, while $\max\left\{\delta_1^{H,tacit}, \delta_2^{H,tacit}\right\} = \max\left\{0.603, 0.688\right\} = 0.745$.

With this graphical analysis, we have proved that, for all $\alpha_1, \alpha_2 \in [0, 1]^2$ (with $\alpha_1 \leqslant \alpha_2$) and for all $c \in (0, 1)^2$, $\delta^{expl} \geqslant \delta^{tacit}$. □

Corollary 2.1. *When agents have complementary thresholds (i.e. $\alpha_1 + \alpha_2 \approx 1$), explicit collusion is more profitable than tacit collusion, even though the former is more difficult to sustain than the latter.*

Proof. Again, we prove this Proposition through a graphical analysis. In Figure 2.19, we plot (in red) the locus of points in $\alpha_1, \alpha_2 \in [0,1]^2$ (given $\alpha_1 \leqslant \alpha_2$) where the following system of inequalities is satisfied.

$$\begin{cases} E_i\left[\Pi_i^{expl}\right] \geqslant E_1\left[\Pi_1^{tacit}\right] \\ E_i\left[\Pi_i^{expl}\right] \geqslant E_2\left[\Pi_2^{tacit}\right] \end{cases}$$

where $i = 1, 2$ (as, when they explicitly collude, $E_1\left[\Pi_1^{expl}\right] = E_2\left[\Pi_2^{expl}\right]$ for $\alpha_1, \alpha_2 \in [0,1]^2$ and $\alpha_1 \leqslant \alpha_2$).

We plot the condition for different values of c, i.e. $c = \{0.1, 0.2, 0.3, ..., 0.99\}$ (from up-left to down-right).

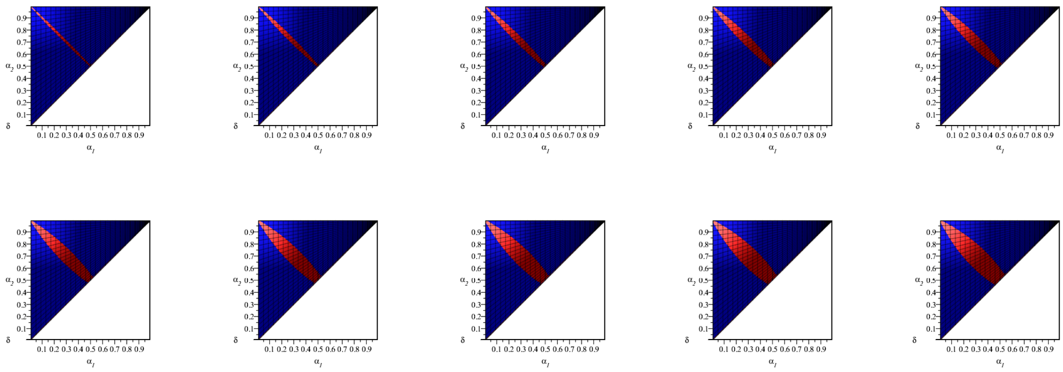


FIGURE 2.19: Conditions under which explicit collusion is more profitable than tacit collusion (in red).

Hence, we have shown that, under certain conditions, explicit collusion trumps tacit collusion. These conditions namely regard the values of the thresholds α_1 and α_2, which need to be nearly complementary, i.e. $\alpha_1 + \alpha_2 \approx 1$. □

Bibliography

[Athey and Bagwell, 2001] Athey, S. and Bagwell, K., 2001. “Optimal collusion with private information”, *The RAND Journal of Economics*, **32** (3), 428–465.

[Bos et al., 2018] Bos, I., Davies, S., Harrington, J.E., and Ormosi, P.L., 2018. “Does enforcement deter cartels? A tale of two tails”, *International Journal of Industrial Organization*, **59**, 372–405.

[Calvano et al., 2019] Calvano, E., Calzolari, G., Denicoló, V., and Pastorello, S., 2019. “Artificial intelligence, algorithmic pricing and collusion”, Available at SSRN: https://ssrn.com/abstract=3304991.

[Compte et al., 2012] Compte, O., Jenny, F., and Rey, P., 2002. “Capacity constraints, mergers and collusion”, *European Economic Review*, **46** (1), 1–29.

[Deng, 2018] Deng, A., 2018. “What do we know about algorithmic tacit collusion”, *Antitrust*, **33** (1), Fall 2018, 88–95.

[Davies et al., 2011] Davies, S., Olczak, M., and Coles, H., 2011. “Tacit collusion, firm asymmetries and numbers: Evidence from EC merger cases”, *International Journal of Industrial Organization*, **29** (2), 221–231.

[Ezrachi and Stucke, 2016] Ezrachi, A., and Stucke, M.E., 2016. “Virtual competition”, *Journal of European Competition Law & Practice*, **7** (9), 585–586.

[Fonseca and Normann, 2012] Fonseca, M.A., and Normann, H.-T., 2012. “Explicit vs tacit collusion – The impact of communication in oligopoly experiments”, *European Economic Review*, **56** (8), 1759–1772.

[Garrod and Olczak, 2017] Garrod, L. and Olczak, M., 2017. “Collusion under imperfect monitoring with asymmetric firms”, *The Journal of Industrial Economics*, **65** (3), 654–682.

[Garrod and Olczak, 2018] Garrod, L. and Olczak, M., 2018. “Explicit vs tacit collusion: The effects of firm numbers and asymmetries”, *International Journal of Industrial Organization*, **56** (1), 1–25.

[Green and Porter, 1984] Green, E.J. and Porter, R.H., 1984. “Noncooperative collusion under imperfect price information”, *Econometrica*, **52** (1), 87–100.

[Harrington, 2008] Harrington, J.E., 2008. “Detecting cartels”, in Buccirossi, P. (ed.), *Handbook of Antitrust Economics*, The MIT Press, 213–258.

[Harrington, 2017] Harrington, J.E., 2017. *The Theory of Collusion and Competition Policy*, The MIT Press.

[Hay and Kelley, 1974] Hay, G. and Kelley, D., 1974. “An empirical survey of price fixing conspiracies”, *The Journal of Law and Economics*, **17** (1), 13–38.

[Levenstein and Suslow, 2006] Levenstein, M.C. and Suslow, V.Y., 2006. “What determines cartel success?”, *Journal of Economic Literature*, **44** (1), 43–95.

[Motta, 2004] Motta, M. ed., 2004. *Competition Policy. Theory and Practice*, Cambridge University Press.

[Obara and Zincenko, 2017] Obara, I. and Zincenko, F., 2017.“Collusion and heterogeneity of firms”, *The RAND Journal of Economics*, **48** (1), 230–249.

[Rotemberg and Saloner, 1986] Rotemberg, J.J. and Saloner, G., 1986. “A supergame-theoretic model of price wars during booms”, *The American Economic Review*, **76** (3), 390–407.

[Spagnolo, 2008] Spagnolo, G., 2008. “Leniency and whistleblowers in Antitrust”, in Buccirossi, P. (ed.), *Handbook of Antitrust Economics*, The MIT Press, 259–303.

[Vives, 2009] Vives, X. ed., 2009. *Competition Policy in the EU*, Oxford University Press.

[Waichman et al., 2014] Waichman, I., Requate, T., and Siang, C.K., 2014. “Communication in Cournot competition: An experimental study”, *Journal of Economic Psychology*, **42**, 1–16.

Chapter 3

Categorical Communication and Team Inefficiencies

Abstract

This work investigates two channels of inefficiencies that help explain the heterogeneity in performance of teams with ex ante similar conditions. We present a simple model of team working where agents need to decide whether to implement or not a potentially profitable project: only the correct choice yield positive profits. We analyze a team composed of two agents. In order to improve the probability to make the right choice, agents may decide to communicate. We assume agents have cognitive limitations: they categorize and are compelled to redefine their communication at a coarser level; hence, inefficiencies may rise due to categorical communication. We find that categorical communication always decreases expected profits compared to perfect communication, but there exist types of agents for which these (inevitable) inefficiencies are minimized. We also study the effects of biased beliefs of team members: not only these beliefs may induce agents to underestimate the profitability of the project, but they may even make communication between team members detrimental.

Keywords: Team working, Communication, Categorization, Biased expectations.

JEL Classification Numbers: D23, D83.

3.1 Introduction

Working in team has a number of positive features, such as the possibility to share knowledge and ideas, wins and losses. The aim of an organization is to coordinate decisions and actions, motivating team members in performing their activities ([Gibbons and Henderson, 2013]). When a team operates efficiently, synergies might rise. Nonetheless, negative consequences may also emerge. In particular, when team members communicate with each others there is room for misunderstandings, and this can lead to inefficiencies.

The purpose of this chapter is to study two types of inefficiencies that arise in team working. Firstly, we investigate how team members' inefficient communication affects the expected profit of the team. Secondly, we study team members' biased beliefs and their impact on team's expected profits.

We analyze the case of a dyad (i.e. a team composed of two members) where each member is expert in a specific field. The first analyzed source of inefficiency arises when the two agents communicate: in particular, as they have different expertise, agents are compelled to redefine their messages at a coarser level. The second source of inefficiency derives from agents' bias in their personal beliefs regarding the profitability of team's production.

To show the impact of these factors on team performance, we consider a simple model where we address the problem of the two teammates who need to decide whether to implement (or discard) a new project whose profitability depends on project's imperfectly observed value and the average value of similar projects in the market (for instance, the prototype of a new product). Consider the following example: a start-up company has developed a new phone-prototype, and its team members have to decide whether to launch it on the market or not. The decision is risky: if the phone is very good and they decide to launch it, it will be a certain gain. However, if they launch it but people do not like it, then it will be a secure loss. The phone is characterized by two dimensions (software and hardware) and each team member (a programmer and an engineer) is expert in one of these.

The value of the project is given by the value of the two dimensions it includes, where each dimension represents a specific feature of the project (e.g. the hardware components). When they evaluate a project, each team member privately observes the value of only one dimension, the one in which she is expert. Moreover, she cannot precisely communicate the observed value because the other agent would not understand it. In other words, as in [Wernerfelt, 2004], agents are compelled to coarsen their messages in order to be understood by the teammate: we call it *categorical communication*.

We consider two types of dyad: *horizontally* or *vertically* structured. In horizontal dyads, both agents have decision power in terms of implementing or rejecting the project, whereas in vertical teams only one agent has. If the team decides to implement the project, and the project value is above the average market value, then the project is profitable and the team earns a positive payoff. On the contrary, when the team implements a project whose value is below the average, they suffer a negative payoff.[1] In order to have more information before deciding, agents may exchange messages regarding the value of the dimension each of them privately observes. Still, because agents only observe the value of one dimensions and the messages regarding the other dimension are coarse, each agent needs to rely on expectations. As expected, our analysis points out that categorical communication decreases the dyads' expected payoffs with respect to perfect

[1] Similarly, the team earns a positive (negative) payoff when it rejects a project whose value is below (above) average.

communication. However, we find conditions under which decision makers can mitigate these inefficiencies: for instance, if they structure the dyad vertically. In fact, under specific conditions it is efficient to centralize information in the hands of only one agent and let her be the receiver of the communication: pyramidal organizations tend to be more efficient than nonpyramidal ones, as in [Bolton et al., 2013].

As an additional source of potential inefficiency, we then assume that agents have biased beliefs: in fact, they may unconsciously consider the project more (or even less) profitable than it actually is. In our work, we assume that this bias induces agents to underestimate the profitability of the project: they reject more projects than it would be correct to do, hence increasing the probability to discard profitable projects. Clearly, we find that such bias is always detrimental to team's expected profits. The central point of the analysis on biased beliefs is that these biases possibly lower the value of the information exchanged: if agents' bias is high enough, the whole team could benefit from lowering the number of messages exchanged. When agents have biased expectations, a pyramidal structure may then not be efficient anymore.

Communication inefficiencies and biased beliefs could help explain why new-born organizations end up either failing or having great success in the market. For instance, as reported by [Everard, 2017], in recent years many Kickstarter projects were unsuccessful because of planning and communicating difficulties among the founders.[2]

The chapter is organized as follows. The rest of this section assesses the related literature. Section 3.2 presents the model and the assumptions. Section 3.3 introduces the first layer of potential inefficiencies, i.e. binary categorization, assuming unbiased beliefs. We analyze expected profits of all possible team structures (vertical and horizontal teams) and communication types (no communication, unilateral and bilateral). In Section 3.4 we introduce biased expectations, and we study their effect on teams' expected profits: we revisit the previous results given agents have biased beliefs. Section 3.5 presents a discussion of the results, and elicits a possible way to mitigate the inefficiencies caused by categorical thinking and biased expectations. In this section we also discuss the assumption of truthful communication. Section 3.6 concludes. Appendix A includes the computations needed to calculate the expected profits of all possible team settings. Appendix B includes the proofs of the propositions stated along the chapter.

3.1.1 Related literature

This work is related to different strands of literature. The first one regards diversity among team members and organizations (see [Gibbons and Roberts, 2012] for a detailed survey). The issue of diversity among team members and how it relates with problem solving has long been investigated (see [Hong and Page, 2001], [LiCalzi and Surucu, 2012], and [Bolton et al., 2013] among the others). There is general consent that diversity tends to outperform homogeneity when the team is involved in tasks that require the achievement of a correct answer, although the contrary is true when the task requires the achievement of consensus. [Brynjolfsson and Milgrom, 2013] have investigated the conditions under which a well managed organization fosters the emergence of synergies and complementaries among team members. This largely depends on how information is shared, and how efficiently team members can communicate when they have different information. Looking at how information is shared among team members, [Prat, 1996]

[2]Kickstarter is an American public-benefit corporation that operates a global crowdfunding platform focused on creativity, and it helps relatively small firms (some of them are even individual businesses) gathering money, donated by backers, to start and implement their projects.

points out that when they are endowed with different information structures (*diversified knowledge*) the team itself can make more precise decisions regarding a wider range of observations and cases. However, by choosing to give to each of its members a common information structure (*shared knowledge*), the organization can foster coordination among collectively taken decisions.

Our work relates also to the literature regarding categorization. As explained in Chapter 1, people tend to group information in categories, and these categories help the human mind to think and interpret whatever information is registered. Recent works (see, for instance, [Mengel, 2012] and [Heller and Winter, 2016]) have developed the idea that basic categories may be the result of enduring dynamic processes: people form their way to process information emulating other members of the society they live in or other people they interact with . This may cause different groups to categorize differently the same piace of information. Another driver of diversity in categorical thinking is the level of expertise. Categories are characterized by different degrees of granularity, which are given by the level of agent's expertise on the specific object or information categorized. In other words, the operation of categorization has a hierarchical structure: people first identify objects at the basic level and then access the superordinate or subordinate categories. [Tanaka and Taylor, 1991] have proved that experts have finer-grained categories than others.[3] Inspired by this idea, we let agents' diversity be driven by their level of categorization over each dimension: fine-grained categorization for the dimension in which they are expert, coarse-grained for the other one. Consistently with [Gary and Wood, 2010], we find that coarser-grained categories imply lower team-performance outcomes.

The last strand of literature which we refer to is that of biased beliefs of decision makers. Agents make decisions depending on their beliefs, but these beliefs do not always coincide with the truth. Numerous works have extensively investigated the role of biased beliefs as a driver of agents' decisions, especially through models that explain how agents update their biased priors depending on feedbacks. For instance, [Rotemberg and Saloner, 1993] and [Van Den Steen, 2005] have analyzed how leaders' beliefs may impact on the performance of a firm. The former has found conditions under which an organization may benefit from having different types of leaders; the latter, that strong beliefs increase the overall variance of teams' performance. Another important aspect investigated in this literature is the impact of overconfidence. For instance, [Hvide, 2002] finds conditions under which overconfidence (or underconfidence) may emerge in the long run, given that agents take into account *what pays* rather than *what is true*: an overconfident agent assesses his ability to be higher than its true value. On the contrary, [Camerer and Lovallo, 1999] and [Koellinger et al., 2007] find conditions under which overconfidence does not pay off. We depart from the definition of overconfidence as a belief about personal skills, and we interpret it as an overestimation of the average values of dimensions and projects in the market: in order not to generate confusion, we neutrally label it as *bias*. The interpretation of the bias is similar to that of $overconfidence_2$ presented in [Hvide, 2002]: it relates to agents that, having an altered distribution of posteriors, are too certain about some event. In our setting, they unconsciously underestimate the probability that the project's value is above the average market one.

[3] In their experiment, [Tanaka and Taylor, 1991] found that, for instance, the expert birdwatcher knows more than the novice with respect to the characteristics of specific kind of birds (e.g. robins, sparrows) but may not necessarily be more knowledgeable about the general characteristics that distinguish birds from other kinds of animals. In another experiment, [Li et al., 1998] find that white basketball fans are better with respect to other white people in recognizing faces of black people: they developed this "ability" watching a large amount of american basketball games (where the majority of players are black).

3.2 The model

Consider two agents, labeled 1 (male) and 2 (female), that work together and form a dyad. The scope of the dyad is to decide whether a single *project* is worth to be developed or not. We consider a project to be a costless piece of production that is possibly profitable (for instance, a new product or prototype). The profitability condition depends on the *value* of the project, i.e. a numerical attribute that defines the project, and this value hinges on the value of the *dimensions* that the project includes: a dimension is a valuable feature of the project. As there are two team members, we assume that projects include only two dimensions.[4] Each agent has incomplete information because she only observes the value of one dimension, so, when deciding whether to implement the project or not, agents need to rely on expectations.

Dimensions' value. We denote X_i the value of each dimension $i = 1, 2$, and we let X_i be a random variable independently uniformly distributed on the support $R_X = (0, 1)$ for both dimensions (hence, R_X is a totally ordered set). With a slight abuse of notation, we let index i characterize the agent who privately observes the value of dimension i of the project.

Project's value. The value of a project is the sum of the values of the dimensions that it includes: the greater the realization of each dimension, the greater the value of the project.[5] We let Θ be the random variable that describes the value of the project. We have that

$$\Theta = X_1 + X_2$$

A project is *successful* (or *good*) whenever its value is greater than or equal to the average value of the project (which, as dimensions' values are uniformly distributed, is unitary), i.e. if

$$\theta \geqslant E[\Theta] = 1$$

Otherwise, it is *unsuccessful* (or *bad*). Equivalently, we can interpret $E[\Theta]$ as the expected value of the projects in the market. Intuitively, successful project yield a positive payoff if implemented, and a similar profit is earned if an unsuccessful project is discarded (resembling a foregone loss). On the contrary, an unsuccessful project that is implemented and a successful project that is discarded yield a negative payoffs. Hence, the team aims at implementing successful projects and discard unsuccessful ones. The ultimate decision whether to implement or discard a project depends on the individual decision of each team member.

Voting rule. Let us denote v_i the vote expressed by agent $i = 1, 2$ to manifest her individual decision: $v_i = A$ ($v_i = R$) means that agent i wants to implement (reject) the project. In order to maximize the probability to obtain a positive payoff, agents vote

[4]For instance, Bic pens, lighters and razors are simple products because they involve only two dimensions: plastic injection molding and marketing. This example is taken from the presentation that Luis Garicano gave at the MIT Seminar in Organizational Economics, April 25th 2017, about "The economics of corporate strategy: How firms leverage their capabilities to make diversification decisions".

[5]As in [Alfaro et al., 2018].

according to the following condition:

$$v_i = \begin{cases} A & \text{iff } x_i + E_i\Big[X_j\Big] \geqslant E_i\Big[\Theta\Big] \\ R & \text{otherwise} \end{cases}$$

In other words, agent i prefers to implement the project if the sum of the observed value x_i of her dimension i and the expected value $E_i\Big[X_j\Big]$ of the unknown dimension j is higher than the expected value of the project $E_i\Big[\Theta\Big]$, where $i = 1, 2$ and $j = 3 - i$. From this condition we can surmise a *voting rule*, i.e. a decision criterion that determines agent i's vote:

$$v_i = \begin{cases} A & \text{iff } x_i \geqslant \bar{x}_i \\ R & \text{otherwise} \end{cases} \tag{3.1}$$

where $\bar{x}_i = E_i\Big[\Theta\Big] - E_i\Big[X_j\Big]$ is the limit value of x_i above which agent i votes A and below which she votes R.

Decision. The ultimate decision of implementing or rejecting a project may involve both agents or just one of them, depending on the structure of the dyad. In a dyad with *horizontal structure*, both team members' vote matter. For instance, if they both want to discard the project, the project is discarded. Because disagreements are possible (e.g. $v_1 = A$ and $v_2 = R$), as a tie breaking rule we assume that projects are implemented with probability $\frac{1}{2}$ when team members' votes are different. In a team with *vertical structure* the ultimate decision is made by just one agent (called *leader*), hence disagreements cannot emerge. The structure is exogenously assigned *ex ante*, and we assume agents cannot modify it.[6]

Let $\bar{v}$ denote the dyad's ultimate decision of implementing or rejecting a project. We have that in a vertical structure $\bar{v} = v_i$ (where i is the leader). If the dyad is horizontally structured, then two possible cases arise: (i) unanimity, and thus $\bar{v} = v_i = v_j$, and (ii) disagreement, where $\bar{v} = \Big\{\frac{1}{2}A, \frac{1}{2}R\Big\}$.

Payoffs. Let $\pi(\bar{v}, \theta)$ denote the *ex post* payoff (hereafter, also *profits* or *performance*). Payoff depends on the realized project's value and the final decision $\bar{v}$. As mentioned before, positive payoff is realized when a successful project is implemented or an unsuccessful one is discarded. Conversely, a negative payoff arise whenever an unsuccessful project is implemented or a successful project is discarded. More precisely, we have

$$\begin{cases} \pi(\bar{v} = A|\theta \geqslant 1) = +1 \\ \pi(\bar{v} = R|\theta < 1) = +1 \\ \pi(\bar{v} = A|\theta < 1) = -1 \\ \pi(\bar{v} = R|\theta \geqslant 1) = -1 \end{cases}$$

[6]This assumption resembles the case of a team with a precise hierarchical structure that cannot be changed in the short run.

After the game has ended and payoff is realized, all ambiguities are solved: agents observe both dimensions' value and, hence, the project's value. This paper adopts a team-theoretical framework, so agents share the total output by equal proportions.

Example 3.2.1. *Let $x_1 = 0.3$ and $x_2 = 0.8$ be the values of the hardware and software dimensions respectively. Hence, the project is successful. According to rule (3.1), agent 1 votes $v_1 = R$ because $x_1 < \bar{x}_1 = 0.5$, while agent 2 votes $v_2 = A$ because $x_2 > \bar{x}_2 = 0.5$. If the team has a vertical structure, they earn $+1$ if agent 2 is the leader and -1 if agent 1 is. If the structure is horizontal, they flip a coin and earn 0 on average.* △

Agents' expertise. Each agent only observes the value of one dimension, the one in which she is *expert*. With respect to the unobserved dimension, agents only have coarse cognition: we assume they can only discriminate two categories of the value of the unobserved dimension. Consistently with the previous chapters, we let a categorization C_i be a partition of R_X. Agent i categorizes the support of dimension j in two convex categories (i.e. intervals), such that $C_i = \{C_i^L, C_i^H\}$. We assume $C_i^L = (0, \alpha_i)$ and $C_i^H = (\alpha_i, 1)$, where $\alpha_i \in [0, 1]$ is the exogenous and fixed *threshold* of agent i, and labels L and H denote *Low* and *High* values of X_j. For instance, agent 1 cannot perfectly discriminate the value of x_2, and he only recognizes two categories of low and high values.

Communication. As their vote is influenced also by the expected value of the unobserved dimension, it may be useful to have communication between agents so that they can update their expectations about the unobserved dimension. Because agent j does not observe x_i, it would be task of i (who, on the contrary, perfectly observes x_i) to communicate to j information regarding x_i and the pertinent category of C_j which x_i belongs to. Agent i cannot perfectly report the value of x_i to agent j because the latter has a coarse categorization over dimension X_i and can only distinguish two possible categories. Hence, agents are forced to adapt their communication language and categorize the information they transmit depending on each others' threshold. For this to be possible, we assume thresholds are common knowledge.

At the beginning of the game, nature determines the values x_i and x_j as realizations of X_i and X_j respectively. Agent i privately observes x_i and identifies the teammate's category of C_j in which x_i falls. We call $S_j^{x_i}$ the element of C_j to which x_i belongs. For instance, if $x_i \in C_j^L$, we have that $S_j^{x_i} = S_j^L = C_j^L$.

After identifying the pertinent category $S_j^{x_i}$, agent i may decide to send a message to reveal to agent j this piece of information. Let $M_i = \{M_i^L, M_i^H, \emptyset\}$ be the set of possible messages sent by agent i. A non-void message sent by i is an element of the set of categories of j; that is, if $M_i \neq \emptyset$, then $M_i \in C_j$. We assume messages are truthful because they contain the true value: $M_i = S_j^{x_i}$. This information is verifiable after the game has ended: agents can recognize if the sender was being truthful or not, and we assume that there exists a punishment that discourages any sender from being untruthful. This assumption is later discussed in Section 3.5.1.

Communicating via categories is the only way in which agents can understand messages and update their prior beliefs about the value of the unobserved dimension.

After receiving a message, agent i refines her understanding of the value of the unobserved dimension. Hence, after receiving a message M_j, agent i casts her vote depending on the following rule:

$$v_i = \begin{cases} A & \text{iff } x_i \geqslant \bar{x}_i^k \\ R & \text{otherwise} \end{cases} \tag{3.2}$$

where $\bar{x}_i^k = E_i\Big[\Theta\Big] - E_i\Big[X_j | M_j = M_j^k\Big]$ and $k = \{L, H\}$. Rule (3.2) represents the update of rule (3.1) after receiving a message. Communication may be bilateral or unilateral. In bilateral, both agents send messages; in unilateral, only one does.

Example 3.2.2. *Let $x_1 = 0.3$ and $x_2 = 0.8$. Let $\alpha_1 = 0.6$ and $\alpha_2 = 0.45$ be the thresholds of agents 1 and 2 respectively. On the one hand, as $x_2 \in C_1^H = (\alpha_1, 1)$, we have that $S_1^{x_2} = C_1^H$ and, hence, $M_2 = M_2^H$. According to rule (3.2), after this message, because $x_1 > \bar{x}_1^H = 1 - 0.8 = 0.2$, agent 1 votes $v_1 = A$. On the other hand, as $x_1 \in C_2^L = (0, \alpha_2)$, we have that $S_2^{x_1} = C_2^L$ and hence $M_1 = M_1^L$. Thus, agent 2 votes $v_2 = A$, as $x_2 > \bar{x}_2^L = 1 - 0.225 = 0.775$. Both agents vote to implement the project: regardless of their structure, the team earns $+1$.* △

Biased expectations. We introduce personal biases that deviate the individual expectation from the true average of dimensions' value X. As X is uniformly distributed, $F(x) = x$ is the true probability distribution, and we let

$$F_i(x|\mu_i) = x^{1+\mu_i}$$

denote the biased probability distribution perceived by agent i. The parameter $\mu_i \in [0, 1]$ represents the bias. As the bias increases, it pushes the *perceived distribution* away from the uniform case. In fact, when $\mu_i > 0$ then $E_i[X|\mu_i] = \frac{1+\mu_i}{2+\mu_i}$. The bias induces i to overestimate the average value of both dimensions, and hence the average value of the project. We assume biases are equal across dimensions but idiosyncratic to each agent. We take biases as exogenous and we make no assumptions on the agents' awareness about their bias.

If we allow agents to have biased expectations, the voting rule changes as follows:

$$v_i = \begin{cases} A & \text{iff } x_i \geqslant \bar{x}_i^{k,\mu_i} \\ R & \text{otherwise} \end{cases} \tag{3.3}$$

where $\bar{x}_i^{k,\mu_i} = E_i\Big[\Theta|\mu_i\Big] - E_i\Big[X_j | M_i = M_j^k, \mu_i\Big]$ with $k = \{L, H\}$. The superscript μ_i in $\bar{x}_i^{k,\mu_i}$ indicates that agent i has bias μ_i. Rule (3.3) represents the update of rule (3.2) when we let agents have biased beliefs over the distribution of dimensions' values. When $\mu = 0$, the voting rules (3.2) and (3.3) coincide. In the following sections, we analyze both cases of unbiased ($\mu_i = 0$) and biased ($\mu_i > 0$) beliefs.

Timing. The timing of the model is the following:

1. agents are exogenously assigned a team structure (horizontal or vertical);
2. nature draws x_1 and x_2 as realizations of X_1 and X_2;
3. each agent i observes x_i, where $i = 1, 2$;

4. agent i may send message $M_i = S_j^{x_i}$ to agent j. If both agents send messages, they do it simultaneously;
5. if the dyad is horizontally structured, both agents vote simultaneously to implement or reject the project. Otherwise, only the leader votes;
6. the ultimate decision is made, and the project is either implemented or discarded. Payoffs are realized.

Benchmarks. As benchmark cases (we assume unbiased beliefs), let us analyze the game when (i) both agents can only distinguish one category for both dimensions (*no discrimination*), and (ii) they can distinguish every possible value of both dimensions (*perfect discrimination*). Under no discrimination, both agents decide randomly whether to implement the project or not: they will be correct half of the times, and wrong the other half, hence making null expected profits. Under perfect discrimination, they will always make the correct decision, hence making profits equal to one, which is the maximum possible.

As dimensions' values are uniformly distributed, realizations can be represented on the square space $(0,1)^2$. Each point in the square represents a project, and the diagonal divides successful projects (in the north-east corner) from unsuccessful ones (in the south-west corner). Benchmark cases are depicted in Figure 3.1. On the left, no discrimination: they accept or reject all projects with same probability, both the good ones (above diagonal) and the bad ones (below diagonal). On the right, perfect discrimination: they only make the correct decisions, and earn the highest possible payoff.

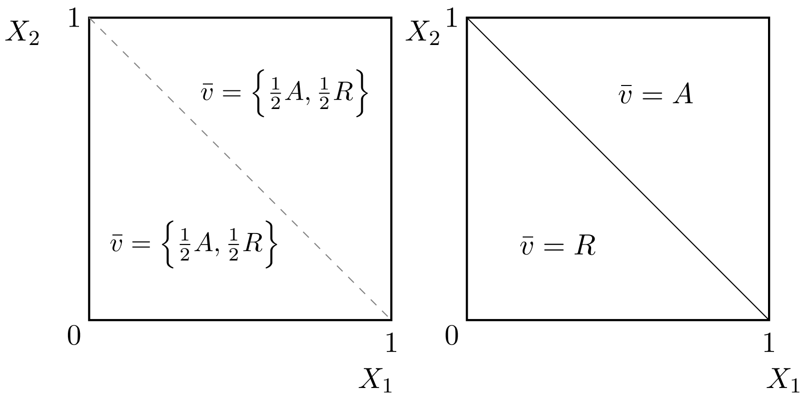


FIGURE 3.1: On the left, agents choose randomly. On the right, agents with perfect categorization.

As each agent has a coarse cognition of the unobserved dimension, and team members' communication is bounded by their categorizations, we find ourselves in a case in between the two described in Figure 3.1. Binary categorization over the unobserved dimension drives agents to incur two different types of errors. Borrowing notation from statistics, we call *type-I error* the act of voting to reject a successful project, and *type-II error* the act of voting to implement an unsuccessful project.

3.3 Categorical rationality

To study the overall effects of categorical communication, we first identify the possible settings of a dyad. A setting is a combination of structure and communication type. In particular, we analyze the two possible structures (vertical and horizontal) and, for each structure, all possible communication types (no communication, unilateral communication, and bilateral communication). Hence, there are in total six possible settings; for each of these settings, we compute the expected profits. Once we have computed the expected profits for all possible settings, we investigate the overall effect of categorical communication. As expected, we find that categorical communication decreases the expected payoff with respect to the case of perfect categorization. Then, among binary categorizers, we find the ones who can guarantee better payoffs to their team. In order to focus on the effects of categorical communication, we let $\mu_i = 0$ for each agent $i = 1, 2$ and discuss the effects of biased beliefs in the next section.

3.3.1 Vertical structure

We first analyze a vertically structured dyad: the only vote that matters for the final decision is the vote of the leader. We separately investigate the cases of (i) no communication, and (ii) unilateral communication. In this latter case, we let the leader be the receiver of the message. The settings of unilateral communication where the sender is the leader and bilateral communication are omitted as they resemble (i) and (ii) respectively.

No communication

Assume, without loss of generality, that agent 1 is the leader. Under no communication, the vote of the leader is not influenced by any message, hence he casts his vote as by condition in (3.1): the leader votes to implement the project only if the value of the observed dimension is greater than its average, i.e. if $x_1 \geqslant \bar{x}_1 = \frac{1}{2}$. Otherwise, he votes to reject it (as the dyad is vertically structured, the final decision $\bar{v}$ coincides with v_1). Voting decision of the leader under no communication can be represented as in Figure 3.2: accept all the projects on the right of $x_1 = \frac{1}{2}$, and reject all those on the left.

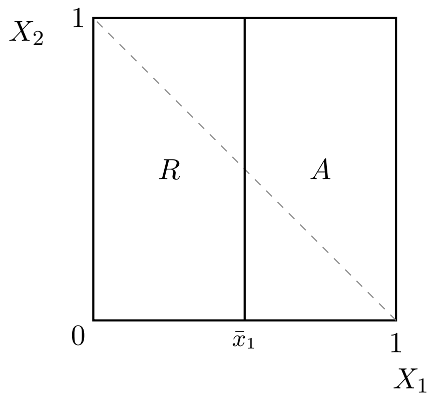


Figure 3.2: Vertical dyad and no communication.

He makes both types of errors with positive probability: he rejects some good projects (the ones where dimension x_1 is low but x_2 is high enough), making the type-I error, and accepts some bad ones (the ones where x_1 is high but x_2 is low enough), making the type-II error.

Let $E\left[\Pi^{V,N}\right]$ denote the *ex ante* expected payoff of the dyad under vertical structure and no communication. To compute it, we sum the *ex post* joint profits for each of the four areas in Figure 3.2, weighted by the probability that x_1 and x_2 fall in each area. As proved in Appendix A, we have

$$E\Big[\Pi^{V,N}\Big] = \frac{1}{2} \tag{3.4}$$

Ex ante expected profits are fixed and positive, as the leader makes the right choice three times out of four.

Unilateral communication

Under unilateral communication, v_1 also depends on the message sent by agent 2 (*sender*): she issues a message M_2 that tells agent 1 if x_2 is low or high. The voting rule of agent 1 is described in (3.2) and states that he votes A if $x_1 \geqslant \bar{x}_1^k$, where we have that

$$\bar{x}_1^k = \begin{cases} \bar{x}_1^L = \frac{2-\alpha_1}{2} & \text{if } M_2 = M_2^L \\ \bar{x}_1^H = \frac{1-\alpha_1}{2} & \text{if } M_2 = M_2^H \end{cases}$$

Voting rule in (3.2) of the leader under unilateral communication can be represented as in Figure 3.3.

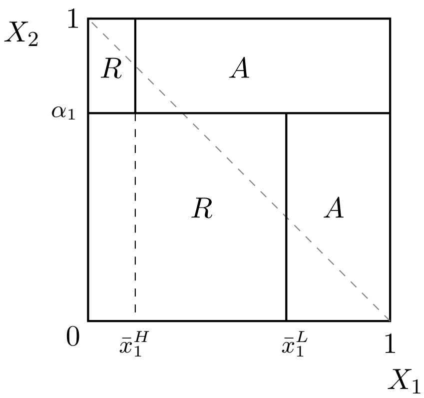


FIGURE 3.3: Vertical dyad and unilateral communication.

Let $E\Big[\Pi^{V,U}(\alpha_i)\Big]$ denote the *ex ante* expected payoff of the dyad under vertical structure and unilateral communication, where agent i is the leader. As proved in Appendix A, the expected payoff is expressed by the following function:

$$E\Big[\Pi^{V,U}(\alpha_i)\Big] = \frac{1}{2} + \alpha_i - \alpha_i^2 \tag{3.5}$$

As a result of agent 2's communication, agent 1 is able to reduce the probability to incur type-I and type-II errors.

3.3.2 Horizontal structure

We now analyze a horizontally structured dyad: the votes of both team members matter for the ultimate decision about the project. We investigate three possible cases: no communication, unilateral communication, and bilateral communication.

No communication

Under no communication, agents' votes are cast only depending on the values of the privately observed dimensions. In fact, each team member $i = 1, 2$ votes to implement the project ($v_i = A$) only if $x_i > \bar{x}_i = \frac{1}{2}$; otherwise she votes to reject it.

Voting decisions of team members are represented in Figure 3.4 and divide projects in four regions, depending on the votes of agents 1 and 2. For instance, "RA" means that $v_1 = R$ and $v_2 = A$ for all the projects included in that region.

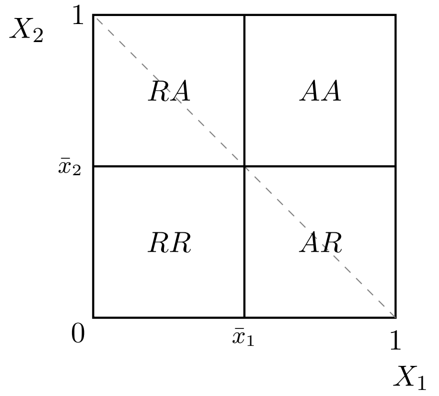


FIGURE 3.4: Horizontal dyad and no communication.

Let $E\left[\Pi^{H,N}\right]$ denote the expected payoff of the team under horizontal structure and no communication. As proved in Appendix A, we have that

$$E\left[\Pi^{H,N}\right] = \frac{1}{2} \tag{3.6}$$

The ultimate decision is always the correct one in the south-west and north-east regions. In the remaining regions, the correct decision is taken only half of the times, hence expected profits are null. Comparing $E\left[\Pi^{V,N}\right]$ and $E\left[\Pi^{H,N}\right]$, found in (3.1) and (3.6) respectively, we see that they are equal, suggesting that having one or two agents with decision power is not different. In fact, the type of structure (vertical or horizontal) does not affect the expected profits when there is no communication among agents. This result is not surprising: without communication, the structure itself may not be clearly defined from an objective perspective, hence involving more team members in the decision process without letting them communicate is not making any difference in terms of expected profits.

Unilateral communication

Under unilateral communication, only the sender issues a message. Suppose agent 1 is the receiver and agent 2 is the sender. The voting rule of agent 1 changes according to the rule described in (3.2), and he votes $v_1 = A$ whenever $x_1 \geqslant \bar{x}_1^L = \frac{2-\alpha_1}{2}$ if $M_2 = M_2^L$, or whenever $x_1 \geqslant \bar{x}_1^H = \frac{1-\alpha_1}{2}$ if $M_2 = M_2^H$. Agent 2's voting rule does not change with respect to the case of no communication: in fact, she does not receive any message to update her expectations with. Hence, she votes $v_2 = A$ when $x_2 \geqslant \bar{x}_2 = \frac{1}{2}$, and $v_2 = R$ otherwise. In Figure 3.5 we identify six different regions of projects for which agent 1 and agent 2 respectively propose their vote.

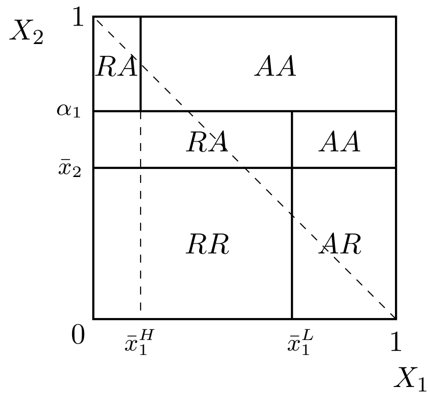


FIGURE 3.5: Horizontal dyad and unilateral communication.

Let $E\Big[\Pi^{H,U}(\alpha_i)\Big]$ denote the *ex ante* expected payoff of the team under horizontal structure and unilateral communication, where agent i is the receiver of the unilateral message. We have that

$$E\Big[\Pi^{H,U}(\alpha_i)\Big] = \frac{1}{2}\Big(1 + \alpha_i - \alpha_i^2\Big) \tag{3.7}$$

Payoff in (3.7) only depends on the threshold of the receiver. This payoff is lower than the payoff of a dyad with an analogous type of communication but vertically structured. This happens because agent 2 votes without having any additional information about x_1, and hence she casts erroneous votes with relatively high probability.

Bilateral communication

We now investigate the case when messages flow in both directions. Team members' voting rule is expressed as in (3.2): whenever $x_i \geqslant \bar{x}_i^k$, agent i votes $v_i = A$, where $\bar{x}_i^k = E_i\Big[\Theta\Big] - E_i\Big[X_j|M_j^k\Big]$ with $i = 1, 2$, $j = 3 - i$, and $k = \{L, H\}$. This case is graphically represented in Figure 3.6, where $\alpha_1 > \alpha_2$ is without loss of generality.

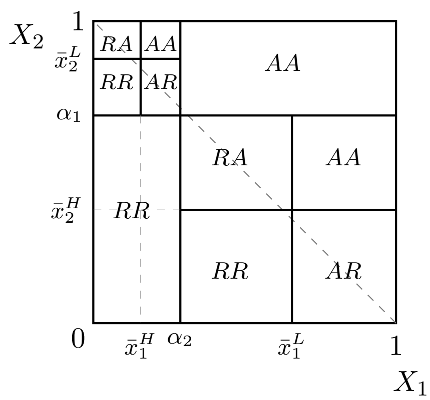


FIGURE 3.6: Horizontal dyad and bilateral communication.

There are ten different regions of projects defined by agents' individual votes. Combining project values, agents' decisions and the probability that a project falls in each region, we can compute $E[\Pi^{H,B}(\alpha_1, \alpha_2)]$, i.e. the *ex ante* expected payoff of a horizontally structured team under bilateral communication. We have that

$$E\Big[\Pi^{H,B}(\alpha_1,\alpha_2)\Big] = \frac{1}{2}\Big(1+\alpha_1-\alpha_1^2+\alpha_2-\alpha_2^2\Big) \tag{3.8}$$

Under bilateral communication, both agents' threshold affects the expected profits, as both agents exchange messages and, moreover, both of them are involved in making the ultimate decision $\bar{v}$.

3.3.3 Effects of categorization and categorical communication

We now summarize the effects of categorical communication on team's expected profits. By analyzing the expected profit functions presented so far, it is possible to infer that, clearly, binary categorization decreases the expected payoff with respect to the case of perfect categorization. This happens because agents have coarse cognition of the value of the dimension in which they are not expert, and hence they need to communicate with coarse messages. With respect to perfect categorization (under which agents can exchange precise messages) categorical communication lowers the expected profits of the dyad: in fact, it increases the probability of making wrong decisions.

Proposition 3.1. *Categorization decreases the expected payoff of the dyad.*

Proof. See Appendix B. □

In fact, independently of the structure and the type of communication, dyad's expected payoff are always lower than 1, i.e. the payoff found for the case of perfect discrimination on page 81. This is not surprising: a coarse discrimination of a dimension yields the agent to lose valuable information. This result is consistent, for instance, with [Gary and Wood, 2010], who find that more accurate mental models (i.e. simplified cognitive representations about how the business environment works) yield to higher performance outcomes.[7]

Given the categorization of team members is binary, we can examine which is the value of the optimal threshold, i.e. which is the optimal way in which agents may categorize.

Proposition 3.2. *Regardless of their structure, dyads where agents communicate (unilaterally or bilaterally) tend to earn higher ex ante expected payoffs when agents' thresholds are closer to 0.5.*

Proof. See Appendix B. □

The proof of this proposition is trivial: all the expected profit functions that depend on α_1, or α_1 and α_2 (i.e. the expected profits of dyads whose agents communicate, unilaterally or bilaterally), exhibit a global maximum in $\alpha_i = 0.5$ (when they depend only on α_i), or $\alpha_1 = \alpha_2 = 0.5$ (when they depend on both thresholds). In fact, note that $E\Big[\Pi^{V,U}(\alpha_i)\Big]$ and $E\Big[\Pi^{H,U}(\alpha_i)\Big]$ are inverted U-shaped parabolas with vertex in $\alpha_i = 0.5$ and global minima in $\alpha_i = \{0,1\}$, while $E\Big[\Pi^{H,B}(\alpha_1,\alpha_2)\Big]$ is an elliptic paraboloid with global maximum in $\alpha_1 = \alpha_2 = 0.5$ and global minima in $\alpha_1 = \{0,1\} \wedge \alpha_2 = \{0,1\}$. In their minima, all these expected profits are equal to $\frac{1}{2}$, exactly as under no communication. In their maximum

[7] The proof of this statement is beyond the scope of this paper, but we believe that in our setting a refinement in agents' categorizations would increase expected profits.

values, $E\Big[\Pi^{H,U}(\alpha_i)\Big]$ is equal to 0.625, while $E\Big[\Pi^{H,B}(\alpha_1, \alpha_2)\Big]$ and $E\Big[\Pi^{V,U}(\alpha_i)\Big]$ are equal to 0.75 (halfway between perfect and no categorization).

This result depends on the assumption of uniform distribution of the random variables X_1 and X_2: for instance, with an asymmetric distribution the optimal threshold would be different. However, it captures the important idea that this threshold's value divides R_X in two intervals of equal length (as 0.5 represents both the median and the mean of X, we cannot tell which moment of the distribution reflects the optimal threshold). Hence, given $\#C_i = 2$, agents with threshold equal to 0.5 are the ones that guarantee the highest expected profits for the team. This happens because these types of agents are able to minimize inefficiencies (i.e. errors), and, on average, make the best decisions.

The difference of expected profits between similar dyads (i.e. dyads with equal structure and communication type) composed of agents with different thresholds is substantial and can be quantified. For instance, consider two vertical dyads. The first one has a leader whose threshold $\alpha = 0.7$: the team expects to gain 0.71. The second one has a leader whose $\alpha = 0.6$, and her dyad's expected profits are 0.74. It is a more than 4% difference, on average. This difference may increase up to 50% if we compare two dyads whose leaders have $\alpha_i = 0.5$ and $\alpha_j = \{0, 1\}$.

3.4 Biased beliefs

Until now, we have assumed agent i's expectation about X coincides with its true average value. Now, inspired by [Hvide, 2002], we introduce a bias in agents' expectations of dimensions' value and, consequently, on the value of projects. We let μ_i denote the parameter that describes agent i's bias. Bias μ_i pushes the perceived distribution away form the true one, attaching higher probabilities to better projects. Some examples of $F_i(x|\mu_i) = x^{1+\mu_i}$ are presented in Figure 3.7: when $\mu_i = 0$, the perceived distribution corresponds with the true distribution; when $\mu_i > 0$, the perceived distribution shifts probability to higher values.

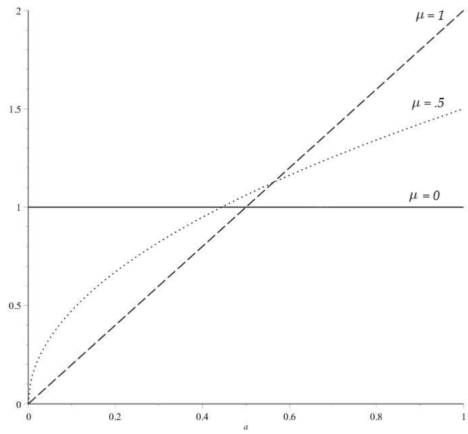


FIGURE 3.7: P.d.f. for different levels of bias μ_i.

The expected value $E_i\Big[X|\mu_i\Big] = \frac{1+\mu_i}{2+\mu_i}$ is increasing in μ_i, meaning that the expected value of the perceived distribution shifts to the right with respect to the expected value of the true distribution. Because $E_i\Big[X|\mu_i > 0\Big] \geqslant E_i\Big[X|\mu_i = 0\Big]$, when voting to implement or discard a project an agent expects the unobserved dimension's value to be higher than it actually is. By this reasoning, he overestimates the overall value of the project. However, we also have that $E_i\Big[\Theta|\mu_i > 0\Big] \geqslant E_i\Big[\Theta|\mu_i = 0\Big]$: the agent overestimates also the average value of similar projects in the market.

Because the true distribution is $F_i(x)$ and not $F_i(x|\mu_i)$, we find that a positive bias is always detrimental to dyad's performance. When she has biased expectations, on the one hand an agent i becomes more selective when voting to implement a project: she will vote *Accept* only if the observed value of dimension i is sufficiently high (and higher than it is necessary). Thus, she makes less type-II errors. On the other hand, the bias induces the agent to turn down projects that should have been implemented: thus, she makes more type-I errors.

We hereafter investigate the role of biased beliefs in modifying team's expected profits. We find that, regardless of the team structure and the type of communication, a bias in expectations always lowers the *ex ante* expected profits. Hence, the negative effects of making more type-I errors overcome the positive effects of incurring less type-II errors.

3.4.1 Vertical structure

Consistently with the previous section, we first analyze a vertically structured dyad (where agent 1 is the leader). As before, we only consider the cases of no communication and unilateral communication.

No communication

When agents do not communicate, the vote of agent 1 only depends on his expectation about the value of the unknown dimension. According to rule (3.1), agent 1 votes to implement the project if $x_1 \geqslant \bar{x}_1^{\mu_1}$, where $\bar{x}_1^{\mu_1} = E_1[\Theta|\mu_1] - E_1[X_2|\mu_1]$. We have that $E_1[\Theta|\mu_1] = \frac{1+\mu_1}{2+\mu_1} + \frac{1+\mu_1}{2+\mu_1}$ and $E_1[X_2|\mu_1] = \frac{1+\mu_1}{2+\mu_1}$, so

$$\bar{x}_1^{\mu_1} = \frac{1+\mu_1}{2+\mu_1}$$

and $\bar{x}_1^{\mu_1} \geqslant \bar{x}_1 = 0.5$ as long as $\mu_1 \geqslant 0$. In Figure 3.8 we can visually interpret the voting rule of the leader agent 1.

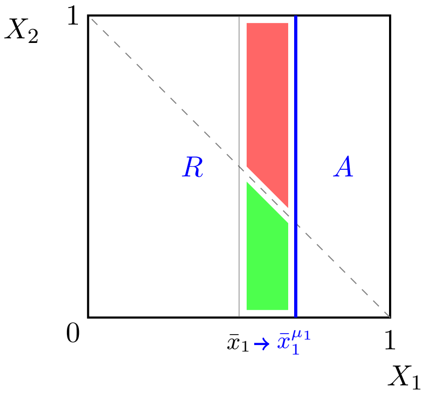


FIGURE 3.8: Vertical dyad with biased leader and no communication: he incurs less type-II errors (green) but more type-I (red) with respect to a unbiased leader.

The blue line represents $\bar{x}_1^{\mu_1}$, and we compare it with the corresponding case in Figure 3.2 where $\mu_1 = 0$. The bias shifts rightward the limit value over which the agent decides to implement the project. It is possible to notice the trade-off expressed above: there is a reduction in the implementation of bad projects (corresponding to the green area), but at the same time there is a more than proportional increase in the rejection of good projects (red area).

Let $E\Big[\Pi^{V,N}(\mu_1)\Big]$ denote the *ex ante* expected payoff of the dyad under vertical structure, no communication, and leader with bias μ_1. We have that

$$E\Big[\Pi^{V,N}(\mu_1)\Big] = \frac{2(1+\mu_1)}{(2+\mu_1)^2}$$

Ex ante expected payoff $E\Big[\Pi^{V,N}(\mu_1)\Big]$ is decreasing in μ_1, proving that the disadvantage from the increase in type-I errors overcomes (in absolute terms) the advantage from the decrease in type-II errors.

Unilateral communication

We now let agent 2 communicate with the leader agent 1, in order to refine the latter's expectation about the unknown dimension. As before, agent 2 reveals to agent 1 if dimension X_2 is low or high, sending the message M_2^L or M_2^H respectively. According to voting rule (3.3), agent 1 prefers to implement the project whenever $x_1 \geqslant \bar{x}_1^{k,\mu_1}$, where $\bar{x}_1^{k,\mu_1} = E_1\Big[\Theta|\mu_1\Big] - E_1\Big[X_2|M_2 = M_2^k, \mu_2\Big]$ with $k = \{L, H\}$. We have that $\bar{x}_1^{k,\mu_1}$ is

$$\bar{x}_1^{k,\mu_1} = \begin{cases} \bar{x}_1^{L,\mu_1} = \frac{2(1+\mu_1)}{2+\mu_1} - \frac{1+\mu_1}{2+\mu_1}\alpha_1 = \frac{(2-\alpha_1)(1+\mu_1)}{2+\mu_1} \text{ if } M_2 = M_2^L \\ \bar{x}_1^{H,\mu_1} = \frac{2(1+\mu_1)}{2+\mu_1} - \frac{1+\mu_1}{2+\mu_1}\frac{1-\alpha_1^{2+\mu_1}}{1-\alpha_1^{1+\mu_1}} = \frac{(1+\mu_1)(1+\alpha_1^{2+\mu_1}-2\alpha_1^{1+\mu_1})}{(2+\mu_1)(1-\alpha_1^{1+\mu_1})} \text{ if } M_2 = M_2^H \end{cases} \tag{3.9}$$

Because $\bar{x}_1^{L,\mu_1} \geqslant \bar{x}_1^L$ and $\bar{x}_1^{H,\mu_1} \geqslant \bar{x}_1^H$, the bias increases both the cutoff values over which the leader prefers to implement the project. We let $E\Big[\Pi^{V,U}(\alpha_1, \mu_1)\Big]$ denote the *ex ante* expected payoff when the team is vertically structured, agent 1 is the leader and he has threshold α_1 and bias μ_1. As the formula is lengthy, we do not report it.[8]

3.4.2 Horizontal structure

When the team is horizontally structured both agents express their vote. We analyze the case of no communication and, then, the other two types of communication (unilateral and bilateral).

No communication

Under no communication, agent $i = 1, 2$ votes $v_i = A$ as long as $x_i \geqslant \bar{x}_i^{\mu_i}$, where $\bar{x}_i^{\mu_i} = \frac{1+\mu_i}{2+\mu_i}$.

We let $E\Big[\Pi^{H,N}(\mu_1, \mu_2)\Big]$ denote the *ex ante* expected payoff of a horizontally structured dyad with biased agents. As proved in Appendix A, this payoff is:

$$E\Big[\Pi^{H,N}(\mu_1, \mu_2)\Big] = \frac{{\mu_1}^2\mu_2 + \mu_1\,{\mu_2}^2 + {\mu_1}^2 + 8\,\mu_2\,\mu_1 + {\mu_2}^2 + 8\,\mu_1 + 8\,\mu_2 + 8}{(\mu_2+2)^2\,(\mu_1+2)^2}$$

[8]We report the general formula in Appendix A.

and only depends on biases μ_1 and μ_2. Agents express the same vote only in two cases: when $x_1 \geqslant \bar{x_1}^{\mu_1}$ and $x_2 \geqslant \bar{x_2}^{\mu_2}$ (they both vote A), or when $x_1 < \bar{x_1}^{\mu_1}$ and $x_2 < \bar{x_2}^{\mu_2}$ (they both vote R). As $\bar{x}_i^{\mu_i} \geqslant \bar{x}_i$, the probability that they vote R increases with respect to the case of zero bias.

Unilateral communication

Under unilateral communication, we assume that only agent 2 sends a message. As she does not receive any message, her voting rule does not change with respect to the previous case: she vote $v_2 = A$ if $x_2 \geqslant \bar{x}_2^{\mu_2} = \frac{1+\mu_2}{2+\mu_2}$, and $v_2 = R$ otherwise.

The receiver agent 1, instead, expresses his vote according to message M_2, exactly as in the case of vertical structure described in (3.9). We let $E\Big[\Pi^{H,U}(\alpha_1, \mu_1, \mu_2)\Big]$ denote the *ex ante* expected payoff of a horizontally structured team where agent 1 is the only receiver of the communication. This payoff depends on the threshold of agent 1 (α_1), his bias (μ_1) and the bias of the other member (μ_2).

Bilateral communication

Under bilateral communication, both agents send and receive messages. Each agent i votes depending on the message M_j of the teammate j according to the following voting rule:

$$\bar{x}_i^{k,\mu_i} = \begin{cases} \bar{x}_i^{L,\mu_i} = \frac{(2-\alpha_i)(1+\mu_i)}{2+\mu_i} & \text{if } \; M_j = M_j^L \\ \bar{x}_i^{H,\mu_i} = \frac{(1+\mu_i)(1+\alpha_i^{2+\mu_i}-2\alpha_i^{1+\mu_i})}{(2+\mu_i)(1-\alpha_i^{1+\mu_i})} & \text{if } \; M_j = M_j^H \end{cases}$$

We let $E\Big[\Pi^{H,B}(\alpha_1, \alpha_2, \mu_1, \mu_2)\Big]$ denote the *ex ante* expected payoff of a horizontally structured team under bilateral communication: it depends on both thresholds and both bias parameters.

3.4.3 Effects of biased expectations

The effect of the bias in agent's expectations is much more straightforward with respect to the effect of categorization: a positive bias always decreases the expected profits. In fact, a positive bias causes the agent to overestimate not only the average value of each dimension, but also the average value of projects. This leads her to turn down projects that should have been implemented: she makes more type-I errors. At the same time, the agent incurs less type-II errors, but the negative effects of the former consequence overcome the positive effects of the latter. Note that we would have found the same result even with underconfident agents, i.e. agents that underestimate the market value of similar projects. In fact, the two cases are symmetric.[9] Hence, we state the following proposition:

Proposition 3.3. *Ex ante expected profits are decreasing in bias μ, independently of dyad's setting.*

Proof. See Appendix B. □

[9] The symmetric case of downward-biased expectations is described by the following distribution function $F_i(x|\mu_i) = (1-x)^{1+\mu_i}$.

A decision maker with $\mu = 0$ is able to minimize overall errors, and hence inefficiencies. Our result is consistent with the work of [Hvide, 2002]. In his dynamic framework, he finds that agents may rationally form overconfident beliefs (called *pragmatic*) that give them a higher expected utility with respect to zero level of overconfidence. This is explained by the fact that the agents take into account *what pays* rather than *what is true*: people tend to forget the failures and remember only the successes. In our model, *what pays* and *what is true* coincide, so agents are better off if they develop beliefs that are as close as possible to the truth. Hence, this work represents an example of how underconfidence (and overconfidence) harm decision makers, and dyads in general.

Value of messages. When agents have unbiased expectations ($\mu_i = 0$ for each $i = 1, 2$), messages always have positive value: regardless of the dyad's structure, expected profits under unilateral or bilateral communication are always greater than profits under no communication. In fact, because $\alpha \in [0, 1]$, messages increase the expected profits of the team because the decision maker(s) has more precise information.[10] This is valid for both vertical and horizontal structures. Under vertical structure, a dyad with unilateral communication performs better than a dyad with no communication.[11] Under horizontal structure, dyads with bilateral communication are better than dyads with unilateral communication, that, in their turn, are better than dyads with no communication. Stated differently, we have that

$$E\Big[\Pi^{V,U}(\alpha_i)\Big] \geqslant E\Big[\Pi^{V,N}\Big] \tag{3.10}$$

and

$$E\Big[\Pi^{H,B}(\alpha_1, \alpha_2)\Big] \geqslant E\Big[\Pi^{H,U}(\alpha_i)\Big] \geqslant E\Big[\Pi^{H,N}\Big] \tag{3.11}$$

for each $\alpha_1, \alpha_2 \in [0, 1]$.

This result profoundly changes if we let agents have biased beliefs. In fact, when agents have biased beliefs ($\mu_i > 0$ for each $i = 1, 2$) rankings in (3.10) and (3.11) do not always hold. It can be useful to interpret messages as suggestions. When the sender of the message has biased beliefs, she would unconsciously transmit her bias to the receiver, issuing a biased suggestion. Hence, the expected profits decrease due to two factors: the bias of the decision maker (the receiver of the message), and the bias of the sender. As a result, we find the paradoxical situation where communication among members may harm the team, especially if members' biases are relatively high. In this case, we say that messages have negative value.

Proposition 3.4. *When agents have biased expectations, messages may have negative value.*

Proof. See Appendix B. □

In Appendix B we find the best possible communication type for vertical and horizontal dyads for each $\alpha_1, \alpha_2 \in (0, 1)^2$ and $\mu_1, \mu_2 > 0$. We find that, depending on pairs of thresholds and biases, exchanging messages may harm the team.

Consider, for instance, a vertical dyad. In Figure 3.9 we plot the following inequality

$$E\Big[\Pi^{V,N}(\mu_1)\Big] \geqslant E\Big[\Pi^{V,U}(\alpha_1, \mu_1)\Big] \tag{3.12}$$

[10] Only in the limit case where sender has threshold $\alpha_i = \{0, 1\}$ her messages have null value, as they do not give informational advantage to the receiver.

[11] Bilateral communication in a vertical dyad represents a case of an increase in the number of messages exchanged but no corresponding increase in the expected profits. However, this does not mean that the message has null value *per se*. It is the fact that the message is not directed to a decision maker that nullifies its value.

as function of α_1 and μ_1. When this inequality is satisfied (grey area), the message sent by agent 2 has negative value.

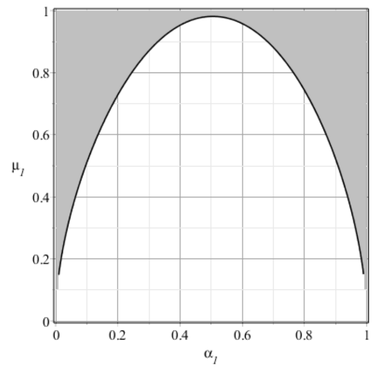


FIGURE 3.9: In grey, the area of (α_1, μ_1) under which condition (3.12) is satisfied.

Messages tend to acquire a negative value the larger is μ_1 and the closer is α_1 to boundary values. The case of horizontal structure is reported in Appendix B, but the overall idea is similar: if μ_i is positive for $i = 1, 2$, messages may harm the dyad, especially if team members' threshold are not central. In the limit case where $\mu_1 = \mu_2 = 1$, there is no region in $\alpha_1, \alpha_2 \in (0,1)^2$ under which exchanging any message is beneficial for the team, regardless of its structure: rankings (3.10) and (3.11) are completely overturned.

If agents are not aware of their bias, they may incorrectly think that having more information may be beneficial. Hence, they may exchange messages and, more importantly, they listen to updating their beliefs, even if this would lead the dyad to a substantial decrease in its expected profits. On the contrary, if agents knew each others' bias, there would exists conditions under which they would rationally choose to discard messages.

3.5 Discussion

The key result of this chapter is that the overall difference in expected profits between dyads that exhibit similar structures (horizontal or vertical) and communication types (unilateral or bilateral) is driven by two parameters: α_i and μ_i, representing respectively the categorization ability and the bias in the beliefs of agent $i = 1, 2$. *Ex ante*, the best possible expected profits are reached when $\alpha_i = 0.5$ and $\mu_i = 0$. Any shift from this case induces a contraction in *ex ante* expected profits. We now analyze whether there exist conditions under which a team composed of agents with certain α and μ can increase their expected payoff.

Throughout this work, which is based on comparative statics between similarly structured dyads, we have assumed that the structure is fixed and exogenous: we have compared dyads without investigating the optimal strategies in terms of, for instance, appointing the leader or deciding how many team members should send messages. If we let agents strategically set their structures and choose their own type of communication, the team could possibly mitigate the inefficiencies.

Suppose agents have zero bias (i.e. $\mu_1 = \mu_2 = 0$). In this case, they would always end up choosing a vertical structure and unilateral communication. An example may help clarify this idea.

Example 3.5.1. *Suppose $\alpha_1 = 0.3$ and $\alpha_2 = 0.6$, and agents have to choose the setting (team structure and communication type) that guarantees the highest ex ante expected* *profits. The strategy of not exchanging messages, regardless of the structure, is dominated by any other choice. If they choose a vertical structure, agents would better appoint agent 2 as leader and agent 1 as (unique) sender, expecting profits equal to* 0.74. *If they choose a horizontal structure, agents would rather bilaterally communicate. In this case, expected profits are* 0.725. △

When agents have no bias and the leader is the agents whose α_i is closest to 0.5, the choice of a vertical structure is always the best one.[12] This result is consistent with [Bolton et al., 2013]. In their work, they introduce communication costs and find that in order to economize these costs it is advantageous to centralize information in the hands of only one agent and let her be the receiver of all communications: an efficient network minimizes the communication channels between agents. The result is that pyramidal organizations are more efficient than nonpyramidal ones, and our work exactly reflects this conclusion.

However, if team members have biased expectations, this unequivocal idea falls. In fact, as previously noted, the ranking between structures and communication types depends (also) on bias parameters.

We have made no assumptions on agents' awareness of their bias. If they are not able to observe it, agents may be tempted to stick to the choice of a vertical structure under unilateral communication. However, if leader's bias is high enough, the advantage of a threshold close to 0.5 could be overcome by the disadvantage of a large bias. On the contrary, if agents are able to observe their own bias, then the choice of structure and communication type depends on four known parameters (thresholds α_1 and α_2 and biases μ_1 and μ_2).[13]

In general, the effects of categorization and bias could potentially be confounded if one were to look only at the performance of the dyad. In fact, the same expected profits could be obtained with different combinations of α and μ. For example, in a vertical dyad with unilateral communication, $E\Big[\Pi^{V,U}(\alpha_1 = 0.5, \mu_1 = 0.4)\Big] = E\Big[\Pi^{V,U}(\alpha_1 = 0.78, \mu_1 = 0)\Big] = 0.67$: equal effects but different causes; the former leader has good categorization ability but a positive bias, whereas the latter has a relatively high threshold but a zero bias. There is no trade-off between α and μ (one parameter cannot compensate the other one), and moreover the effect of μ is additive (because expected profits are decreasing in μ).

The example of Kickstarter, mentioned in the introduction, may provide some empirical evidence: *ceteris paribus*, some of the projects were very successful, whereas recently many projects have declared bankruptcy, "exposing cases where founders' and companies' egos have simply overtaken their ability to reason, plan and communicate logically", as in [Everard, 2017]. Hence, a possible explanation for this phenomenon could be the presence of biased beliefs and team members' inability to communicate efficiently.

The results in this chapter explain differences in expected performance through factors that are endogenous in management practices. There is a wide literature that has provided strong evidence of persistent performance differences among seemingly similar or *ex ante* identical firms, but the question of what causes these differences is still open, as reported in [Gibbons and Henderson, 2013]. In their work, authors find that most of these differences can be related to management practices and managers' abilities. In our work, we find that, net of the type of structure and communication of a dyad, the differences between

[12]Except when $\alpha_1 = \alpha_2 = 0.5$. In this case, agents are indifferent between vertical structure under unilateral communication and horizontal structure under bilateral communication.

[13]Net of preferences for being a leader, agents would always be willing to truthfully communicate their bias to the other team member.

dyads' performances are caused by personal characteristics of the decision maker(s). We have seen that there exist agents with a certain categorization ability and (zero) bias that, on average, make better decisions than the others. In fact, if these agents have a prominent role in the decision making process, the whole dyad would benefit: decision makers' abilities represent the main drivers of dyads' performance. Linking this work to [Gibbons and Henderson, 2013], we find that these factors may contribute in the overall definition of a decision maker's ability and, hence, must be taken into account when addressing the question of the sources of persistent performance differences. In fact, as we have presented them, categories are fixed in the short run and, generally, difficult to shift along time. This may be an explanation why these performance differences tend to endure in the long run.

3.5.1 Truthful agents

In this work we have assumed that messages exchanged by team members are truthful, or, in other words, that the message coincides with the category of agent i's partition to which the true value x_j belongs to. Moreover, this is common knowledge.

Even if the idea that team members communicate truthfully seems reasonable, there exist conditions under which they would find it optimal to send untruthful messages. Consider the following example of a vertical dyad under unilateral communication: agent 1 is the sender and agent 2 is the leader with threshold $\alpha_2 = 0.1$. Suppose that agent 1 observes $x_1 = 0.1 + \varepsilon$, with $\varepsilon > 0$. He has two choices: being truthful and send the *high* message M_1^H, or lie and send the *low* message M_1^L. If agent 1 sends the truthful message, agent 2's decision will be wrong whenever $\bar{x}_2^H = 0.45 < x_2 < 0.9$, because she will *accept* and *unsuccessful* project. However, if agent 1 sends the untruthful message, agent 2's decision will be wrong whenever $0.9 < x_2 < \bar{x}_2^L = 0.95$, because she will *reject* a *successful* project. As the latter situation is better than the former, agent 1 has incentives to lie.

As the focus of this work is on the effects of categorization and biased beliefs and not on the strategic component of messages, we rule out the possibility that agents may lie imposing a strategyproof mechanism that makes truthtelling rational. As noted above, after the game ends all ambiguities are solved, and agents can observe both dimensions' value. This makes messages verifiable: if one agents lies, the other one will know as soon as the game ends. Thus, we imagine a sufficiently large cost born by the agents that sends untruthful messages.

3.6 Conclusions

This chapter studies the impact of categorical thinking and biased beliefs on the expected profits of organized dyads. As widely recognized in cognitive science and economics, agents perceive the information they get from the environment depending on their cognitive limitations. These restraints compel agents to categorize information. In our model, categorical thinking implies categorical communication among team members: agents cannot precisely communicate pieces of information, and they can only send coarse messages. We assume binary categorization and we find two results. The first is that a binary categorization is detrimental with respect to a perfect categorization: incomplete information yield agents commit errors when it is time to take production decisions.[14] The second is that, among binary categorizers, team members that better help teams are the

[14] This result may be generally extended to any coarse categorization.

ones who categorize information in two intervals of equal length: these types of agents minimize the probability to make erroneous decisions.

Then, we investigate the effect of team members' biased expectations on team performance. The literature offers a wide variety of approaches and results. In our model, this bias impacts on the agent's ability to make decisions that match the true state: in particular, it leads them to underestimate the value of the project with respect to the average value of similar projects in the market. The consequence is that agents would be cautious and reject more projects than it would be optimal to do. Hence, we find that any positive bias is always detrimental to the dyad's expected payoff. We also find that, when team members have positive bias, communication may be disadvantageous: we find conditions under which the less the messages exchanged, the higher the expected profits.

Finally, we find actions that agents can do to (possibly) increase dyad's expected profits, such as set the dyad in a vertical structure and appoint as leader the team member with best categorization ability and low bias. However, such actions could be counterproductive if agents have positive bias and they are not aware of it.

This work may be important for empirical analysis where decision makers' abilities, if not adequately measured, can cause an omitted-variable bias. Our results suggest that there exist factors that contribute to establish the social capital of an agent that may be hard to measure of observe. For instance, [Bloom et al., 2012] use trust the proxy social capital and measure its effects on organization and, consequently, on productivity of a firm. Our work suggests that there may exist similar factors that might be relevant in the interpretation of firms' performance, and should be taken into account when measuring it.

Some limiting assumptions imposed along the chapter suggest that this work can be extended in several directions. To begin with, the first natural extension would be to include the possibility of a strategic communication, and let agents send the message they prefer without fearing subsequent punishments. Secondly, we think that augmenting (proportionally) the number of agents and dimensions can increment the robustness of the model and the results. Thirdly (and consistently with the other chapters), our results hinge on the simplifying assumption that the state space is fully ordered, and categories are convex intervals: we believe that relaxing this assumption would add significance to the results. Another important step may be that of allowing for non-uniform distribution of the dimensions: for instance, it would be interesting to use a beta distribution (that is, a generalization of the uniform case). Analyzing non-symmetric distributions may help understanding to which moment of the distribution the optimal threshold coincides. Lastly, we believe it would be interesting to open the model to a dynamic setting, and study how agents may learn to adjust their categorizations and bias, for instance including positive learning costs: we expect to find a trade-off between the benefit of increasing expected profits and the disadvantage of incurring learning costs.

Appendix A

Vertical structure and no communication. We have that

$$\begin{aligned} E\big[\Pi^{V,N}\big] &= \Pr(0 < x_1 < \bar{x}_1) \cdot \gamma_1 + \Pr(\bar{x}_1 < x_1 < 1) \cdot \gamma_2 = \\ &= \frac{1}{2} \end{aligned}$$

where

$$\begin{aligned} \gamma_1 &= \Pr(\theta < 1 | 0 < x_1 < \bar{x}_1) \cdot \pi(\bar{v} = R | \theta < 1) + \Pr(\theta > 1 | 0 < x_1 < \bar{x}_1) \cdot \pi(\bar{v} = R | \theta > 1) = \\ &= 2 \int_0^{\bar{x}_1} \int_0^{1-x} 1 \; \mathrm{d}z \mathrm{d}x \cdot (+1) + 2 \int_0^{\bar{x}_1} \int_{1-x}^{1} 1 \; \mathrm{d}z \mathrm{d}x \cdot (-1) \\ \gamma_2 &= \Pr(\theta > 1 | \bar{x}_1 < x_1 < 1) \cdot \pi(\bar{v} = A | \theta > 1) + \Pr(\theta < 1 | \bar{x}_1 < x_1 < 1) \cdot \pi(\bar{v} = A | \theta < 1) = \\ &= 2 \int_{\bar{x}_1}^{1} \int_{1-x}^{1} 1 \; \mathrm{d}z \mathrm{d}x \cdot (+1) + 2 \int_{\bar{x}_1}^{1} \int_0^{1-x} 1 \; \mathrm{d}z \mathrm{d}x \cdot (-1) \end{aligned}$$

and $\bar{x}_1 = \frac{1}{2}$.

Vertical structure and unilateral communication. Assume without loss of generality that agent 1 is the leader and the receiver of the message.

$$\begin{aligned} E\big[\Pi^{V,U}(\alpha_1)\big] &= \Pr(0 < x_1 < \bar{x}_1^H) \cdot \Pr(\alpha_1 < x_2 < 1) \cdot \gamma_1 + \Pr(\bar{x}_1^H < x_1 < 1) \cdot \Pr(\alpha_1 < x_2 < 1) \cdot \gamma_2 + \\ &\quad + \Pr(0 < x_1 < \bar{x}_1^L) \cdot \Pr(0 < x_2 < \alpha_1) \cdot \gamma_3 + \Pr(\bar{x}_1^L < x_1 < 1) \cdot \Pr(0 < x_2 < \alpha_1) \cdot \gamma_4 = \\ &= \frac{1}{2} + \alpha_1 - \alpha_1^2 \end{aligned}$$

where

$$\begin{aligned} \gamma_1 &= \Pr(\theta < 1 | 0 < x_1 < \bar{x}_1^H, \alpha_1 < x_2 < 1) \cdot \pi(\bar{v} = R | \theta < 1) + \\ &\quad + \Pr(\theta > 1 | 0 < x_1 < \bar{x}_1^H, \alpha_1 < x_2 < 1) \cdot \pi(\bar{v} = R | \theta > 1) = \\ &= \frac{2}{1 - 2\alpha_1 + \alpha_1^2} \left[\int_0^{\bar{x}_1^H} \int_{\alpha_1}^{1-x} 1 \; \mathrm{d}z \mathrm{d}x \cdot (+1) + \int_0^{\bar{x}_1^H} \int_{1-x}^{1} 1 \; \mathrm{d}z \mathrm{d}x \cdot (-1) \right] \\ \gamma_2 &= \Pr(\theta > 1 | \bar{x}_1^H < x_1 < 1, \alpha_1 < x_2 < 1) \cdot \pi(\bar{v} = A | \theta > 1) + \\ &\quad + \Pr(\theta < 1 | \bar{x}_1^H < x_1 < 1, \alpha_1 < x_2 < 1) \cdot \pi(\bar{v} = A | \theta < 1) = \\ &= \frac{2}{1 - \alpha_1^2} \left[\left(\int_{1-\alpha_1}^{1} \int_{\alpha_1}^{1} 1 \; \mathrm{d}z \mathrm{d}x + \int_{\bar{x}_1^H}^{1-\alpha_1} \int_{1-x}^{1} 1 \; \mathrm{d}z \mathrm{d}x \right) \cdot (+1) + \int_{\bar{x}_1^H}^{1-\alpha_1} \int_{\alpha_1}^{1-x} 1 \; \mathrm{d}z \mathrm{d}x \cdot (-1) \right] \\ \gamma_3 &= \Pr(\theta < 1 | 0 < x_1 < \bar{x}_1^L, 0 < x_2 < \alpha_1) \cdot \pi(\bar{v} = R | \theta < 1) + \\ &\quad + \Pr(\theta > 1 | 0 < x_1 < \bar{x}_1^L, 0 < x_2 < \alpha_1) \cdot \pi(\bar{v} = R | \theta > 1) = \\ &= \frac{2}{\alpha_1(2 - \alpha_1)} \left[\left(\int_0^{1-\alpha_1} \int_0^{\alpha_1} 1 \; \mathrm{d}z \mathrm{d}x + \int_{1-\alpha_1}^{\bar{x}_1^L} \int_0^{1-x} 1 \; \mathrm{d}z \mathrm{d}x \right) \cdot (+1) + \right. \\ &\quad \left. + \int_{1-\alpha_1}^{\bar{x}_1^L} \int_{\alpha_1}^{1-x} 1 \; \mathrm{d}y \mathrm{d}x \cdot (-1) \right] \\ \gamma_4 &= \Pr(\theta > 1 | \bar{x}_1^L < x_1 < 1, 0 < x_2 < \alpha_1) \cdot \pi(\bar{v} = A | \theta > 1) + \\ &\quad + \Pr(\theta < 1 | \bar{x}_1^L < x_1 < 1, 0 < x_2 < \alpha_1) \cdot \pi(\bar{v} = A | \theta < 1) = \\ &= \frac{2}{\alpha_1^2} \left[\int_{\bar{x}_1^L}^{1} \int_{1-x}^{\alpha_1} 1 \; \mathrm{d}z \mathrm{d}x \cdot (+1) + \int_{1-\frac{\mu_1}{2}}^{1} \int_0^{1-x} 1 \; \mathrm{d}z \mathrm{d}x \cdot (-1) \right] \end{aligned}$$

and

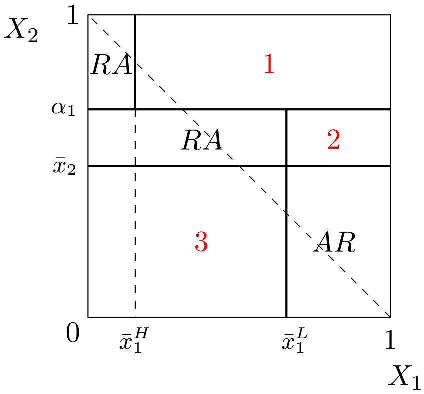
$$\bar{x}_1^L = \frac{2-\alpha_1}{2}$$
$$\bar{x}_1^H = \frac{1-\alpha_1}{2}$$

Horizontal structure and no communication. The expected profits of project included in regions where agents propose different votes have null expected value. Hence, we only consider the regions of projects where agents express the same vote.

$$\begin{aligned} E\left[\Pi^{B,N}\right] &= \Pr(0 < x_1 < \bar{x}_1) \cdot \Pr(0 < x_2 < \bar{x}_2) \cdot \pi(\bar{v} = R|\theta < 1) + \\ &\quad + \Pr(\bar{x}_1 < x_1 < 1) \cdot \Pr(\bar{x}_2 < x_2 < 1) \cdot \pi(\bar{v} = A|\theta > 1) = \\ &= \int_0^{\bar{x}_1} \int_0^{\bar{x}_2} 1 \ \mathrm{d}z\mathrm{d}x \cdot (+1) + \int_{\bar{x}_1}^1 \int_{\bar{x}_2}^1 1 \ \mathrm{d}z\mathrm{d}x \cdot (-1) \\ &= \frac{1}{2} \end{aligned}$$

where $\bar{x}_1 = \bar{x}_2 = \frac{1}{2}$.

Horizontal structure and unilateral communication. The only regions of project where agents express the same vote are the one denoted by a red number in the following figure (agent i is the receiver). Expected profits are expressed by the probabilty that a project belongs to a certain region, multiplied by expected profit of that specific region. Hence we have



$$\begin{aligned} E\left[\Pi^{H,U}(\alpha_1)\right] &= \Pr(\bar{x}_1^H < x_1 < 1) \cdot \Pr(\alpha_1 < x_2 < 1) \cdot \gamma_1 + \\ &\quad + \Pr(\bar{x}_1^L < x_1 < 1) \cdot \Pr(\bar{x}_2 < x_2 < \alpha_1) \cdot \gamma_2 + \\ &\quad + \Pr(0 < x_1 < \bar{x}_1^L) \cdot \Pr(0 < x_2 < \bar{x}_2) \cdot \gamma_3 = \\ &= \frac{1}{2}\left(1 + \alpha_1 - \alpha_1^2\right) \end{aligned}$$

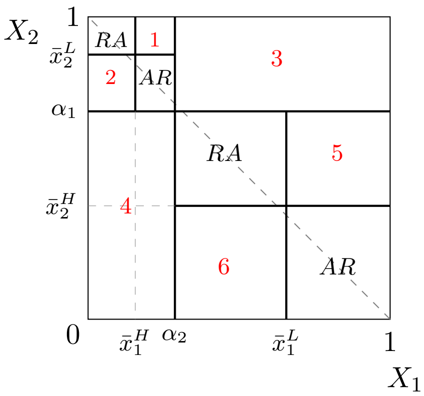
where

$$\begin{aligned} \gamma_1 &= \Pr(\theta > 1|\bar{x}_1^H < x_1 < 1, \alpha_1 < x_2 < 1) \cdot \pi(\bar{v} = A|\theta > 1) + \\ &\quad + \Pr(\theta < 1|\bar{x}_1^H < x_1 < 1, \alpha_1 < x_2 < 1) \cdot \pi(\bar{v} = A|\theta < 1) \\ \gamma_2 &= \Pr(\theta > 1|\bar{x}_1^L < x_1 < 1, \bar{x}_2 < x_2 < \alpha_1) \cdot \pi(\bar{v} = A|\theta > 1) + \\ &\quad + \Pr(\theta < 1|\bar{x}_1^L < x_1 < 1, \bar{x}_2 < x_2 < \alpha_1) \cdot \pi(\bar{v} = A|\theta < 1) \\ \gamma_3 &= \Pr(\theta > 1|0 < x_1 < \bar{x}_1^L, 0 < x_2 < \bar{x}_2) \cdot \pi(\bar{v} = R|\theta > 1) + \\ &\quad + \Pr(\theta < 1|0 < x_1 < \bar{x}_1^L, 0 < x_2 < \bar{x}_2) \cdot \pi(\bar{v} = R|\theta < 1) \end{aligned}$$

and

$$\bar{x}_1^L = \frac{2-\alpha_1}{2}$$
$$\bar{x}_1^H = \frac{1-\alpha_1}{2}$$
$$\bar{x}_2 = \frac{1}{2}$$

Horizontal structure and bilateral communication. As before, the regions denoted by a red number are the ones where agents express the same vote.



Hence we have

$$E\big[\Pi^{H,B}(\alpha_1,\alpha_2)\big] = \Pr(\bar{x}_1^H < x_1 < \alpha_2)\cdot\Pr(\bar{x}_2^L < x_2 < 1)\cdot\gamma_1 + \Pr(0 < x_1 < \bar{x}_1^H)\cdot\Pr(\alpha_1 < x_2 < \bar{x}_2^L)\cdot\gamma_2 + \\ + \Pr(\alpha_2 < x_1 < 1)\cdot\Pr(\alpha_1 < x_2 < 1)\cdot\gamma_3 + \Pr(0 < x_1 < \alpha_2)\cdot\Pr(0 < x_2 < \alpha_1)\cdot\gamma_4 + \\ + \Pr(\bar{x}_1^L < x_1 <_1)\cdot\Pr(\bar{x}_2^H < x_2 < \alpha_1)\cdot\gamma_5 + \Pr(\alpha_2 < x_1 < \bar{x}_1^L)\cdot\Pr(0 < x_2 < \bar{x}_2^H)\cdot\gamma_6$$

where

$$\gamma_1 = \Pr(\theta > 1|\bar{x}_1^H < x_1 < \alpha_2, \bar{x}_2^L < x_2 < 1)\cdot\pi(\bar{v} = A|\theta > 1) + \\ + \Pr(\theta < 1|\bar{x}_1^H < x_1 < \alpha_2, \alpha_1 < x_2 < \bar{x}_2^L)\cdot\pi(\bar{v} = A|\theta < 1)$$
$$\gamma_2 = \Pr(\theta > 1|0 < x_1 < \bar{x}_1^H, \alpha_1 < x_2 < \bar{x}_2^L)\cdot\pi(\bar{v} = R|\theta > 1) + \\ + \Pr(\theta < 1|0 < x_1 < \bar{x}_1^H, \alpha_1 < x_2 < 1)\cdot\pi(\bar{v} = R|\theta < 1)$$
$$\gamma_3 = \Pr(\theta > 1|\alpha_2 < x_1 < 1, \alpha_1 < x_2 < 1)\cdot\pi(\bar{v} = A|\theta > 1) + \\ + \Pr(\theta < 1|\alpha_2 < x_1 < 1, \alpha_1 < x_2 < 1)\cdot\pi(\bar{v} = A|\theta < 1)$$
$$\gamma_4 = \Pr(\theta > 1|0 < x_1 < \alpha_2, 0 < x_2 < \alpha_1)\cdot\pi(\bar{v} = R|\theta > 1) + \\ + \Pr(\theta < 1|0 < x_1 < \alpha_2, 0 < x_2 < \alpha_1)\cdot\pi(\bar{v} = R|\theta < 1)$$
$$\gamma_5 = \Pr(\theta > 1|\bar{x}_1^L < x_1 <_1, \bar{x}_2^H < x_2 < \alpha_1)\cdot\pi(\bar{v} = A|\theta > 1) + \\ + \Pr(\theta < 1|\bar{x}_1^L < x_1 <_1, \bar{x}_2^H < x_2 < \alpha_1)\cdot\pi(\bar{v} = A|\theta < 1)$$
$$\gamma_6 = \Pr(\theta > 1|\alpha_2 < x_1 < \bar{x}_1^L, 0 < x_2 < \bar{x}_2^H)\cdot\pi(\bar{v} = R|\theta > 1) + \\ + \Pr(\theta < 1|\alpha_2 < x_1 < \bar{x}_1^L, 0 < x_2 < \bar{x}_2^H)\cdot\pi(\bar{v} = R|\theta < 1)$$

and

$$\begin{aligned}
\bar{x}_1^L &= \frac{2-\alpha_1}{2} \\
\bar{x}_1^H &= \frac{1-\alpha_1}{2} \\
\bar{x}_2^L &= \frac{2-\alpha_2}{2} \\
\bar{x}_2^H &= \frac{1-\alpha_2}{2}
\end{aligned}$$

Biased expectations. The only difference with the above is in the change of the limit value over which agents decide to vote A. In fact, it is possible to compute expected profits of biased agents just using

$$\begin{aligned}
\bar{x}_1^{\mu_1} &= \frac{1+\mu_1}{2+\mu_1} \\
\bar{x}_2^{\mu_2} &= \frac{1+\mu_2}{2+\mu_2} \\
\bar{x}_1^{L,\mu_1} &= \frac{(2-\alpha_1)(1+\mu_1)}{2+\mu_1} \\
\bar{x}_1^{H,\mu_1} &= \frac{(1+\mu_1)(1+\alpha_1^{2+\mu_1}-2\alpha_1^{1+\mu_1})}{(2+\mu_1)(1-\alpha_1^{1+\mu_1})} \\
\bar{x}_2^{L,\mu_1} &= \frac{(2-\alpha_2)(1+\mu_2)}{2+\mu_2} \\
\bar{x}_2^{H,\mu_1} &= \frac{(1+\mu_2)(1+\alpha_2^{2+\mu_2}-2\alpha_2^{1+\mu_2})}{(2+\mu_2)(1-\alpha_2^{1+\mu_2})}
\end{aligned}$$

instead of $\bar{x}_1$, $\bar{x}_2$, $\bar{x}_1^L$, $\bar{x}_1^H$, $\bar{x}_2^L$, and $\bar{x}_2^H$ respectively.

Appendix B

Proposition 3.1. *Categorization decreases the expected payoff of the dyad.*

Proof. This proof is trivial. When agent has *perfect categorization* over the unknown dimension, the expected profits of the team are 1, the largest possible value. In fact, in this case team members can reciprocally exchange precise messages (i.e. messages that coincide with the privately observed value) and they would be able to perfectly understand them. When we let agents have an imperfect ability to discriminate the unknown dimension (i.e. we let agents have a finite number of categories) we introduce the possibility that they make type-I and type-II errors. These errors are absent when agents have perfect categorization (or, for a vertical team, when the leader has perfect categorization). In fact, as long as $\#C = 2$, threshold α is unique and expected profits are always lower than 1 (regardless of team structure and communication type). □

Proposition 3.2. *Regardless of their structure, dyads where agents communicate (unilaterally or bilaterally) tend to earn higher ex ante expected payoffs when agents' thresholds are closer to 0.5.*

Proof. For a vertical team where agent $i = 1, 2$ is the leader, we have that

$$E\big[\Pi^{V,U}(\alpha_i)\big] = \frac{1}{2} + \alpha_i - \alpha_i^2$$

This is an inverted U-shaped parabola with vertex in $\alpha_i = 0.5$. Hence, expected profits are maximum (and equal to 0.75) when $\alpha_i = 0.5$.

For a horizontal team where only agent i receives messages, we have that

$$E\big[\Pi^{H,U}(\alpha_i)\big] = \frac{1}{2}\Big(1 + \alpha_i - \alpha_i^2\Big)$$

Again, this in another inverted U-shaped parabola with vertex in $\alpha_i = 0.5$. When $\alpha_i = 0.5$, expected profits are 0.625 (half-way between no communication and unilateral communication under vertical stucture).

For a horizontal team where both agents send (and hence receive) messages, we have that expected profits are

$$E\big[\Pi^{H,B}(\alpha_1, \alpha_2)\big] = \frac{1}{2}\Big(1 - \alpha_1^2 + \alpha_1 - \alpha_2^2 + \alpha_2\Big)$$

This function represents a paraboloid with vertex in $(\alpha_1 = 0.5, \alpha_2 = 0.5)$. When $(\alpha_1 = 0.5, \alpha_2 = 0.5)$, expected profits are 0.75. □

Proposition 3.3. *Ex ante expected profits are decreasing in bias μ, independently of dyad's setting.*

Proof. We start from vertical structure under no communication (agent i is the leader), where $E\big[\Pi^{V,N}(\mu_i)\big] = \frac{2(1+\mu_i)}{(2+\mu_i)^2}$. The derivative

$$\frac{\partial E\big[\Pi^{V,N}(\mu_i)\big]}{\partial \mu_i} = -\frac{2\mu_i}{(2+\mu_i)^3}$$

is positive only if $\mu_i \in (-2, 0)$. As we assume $\mu_i \in [0, 1]$ for each $i = 1, 2$, then $E\big[\Pi^{V,N}(\mu_i)\big]$ is decreasing in μ_i.

For the unilateral communication case, the function $E\big[\Pi^{V,U}(\alpha_i, \mu_i)\big]$ is much more complicated, and so it is the computation of the derivative with respect to μ_i. Hence, we prove the statement using different graphs. We plot (in black) the function for different values of $\mu_i = [0.1, 0.2, 0.3, ..., 1]$ and compare it with the case when $\mu_i = 0$ (in grey). The result is plotted in Figure 3.10.

For the case of horizontal dyad, we assume, without loss of generality, that $\mu_1 = \mu_2 = \mu$. This assumption is without loss of generality because both $E[X|\mu_1] = \frac{1+\mu_1}{2+\mu_1}$ and $E[X|\mu_2] = \frac{1+\mu_2}{2+\mu_2}$ linearly increase in μ_1 and μ_2 respectively. Thus, we can better delineate and present the proofs, keeping in mind that results hold also for the case $\mu_1 \neq \mu_2$.

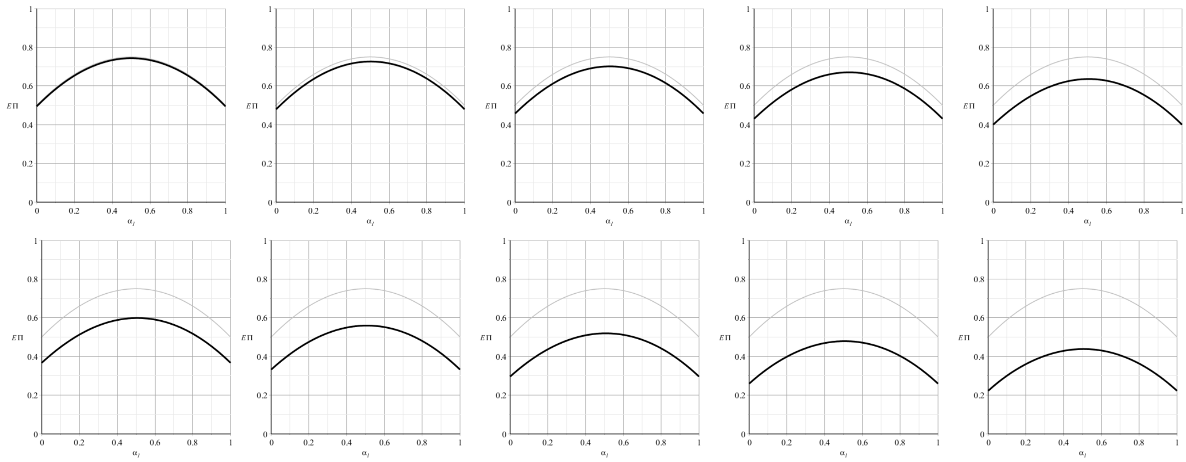


FIGURE 3.10: $E\big[\Pi^{V,U}(\alpha_i,\mu_i)\big]$ for $\mu_i=[0.1,0.2,0.3,...,1]$.

Under no communication, the profit function is

$$E\Big[\Pi^{H,N}(\mu_1,\mu_2)\Big]=\frac{{\mu_1}^2\mu_2+\mu_1\,{\mu_2}^2+{\mu_1}^2+8\,\mu_2\,\mu_1+{\mu_2}^2+8\,\mu_1+8\,\mu_2+8}{\left(2+\mu_2\right)^2\left(2+\mu_1\right)^2}$$

and, if $\mu_1=\mu_2=\mu$, it becomes

$$E\Big[\Pi^{H,N}(\mu)\Big]=\frac{2(1+\mu)}{(2+\mu)^2}$$

resembling the case of vertical structure and no communication.

Under unilateral communication, the profit function $E\big[\Pi^{H,U}(\alpha_1,\mu_1,\mu_2)\big]$ is too long to be reported. Hence, we let $\mu_1=\mu_2=\mu$ and, similarly to the previous case, we plot the function for different values of $\mu=[0.1,0.2,0.3,...,1]$, and compare it with the case $\mu=0$. The graphs are presented in Figure 3.11.

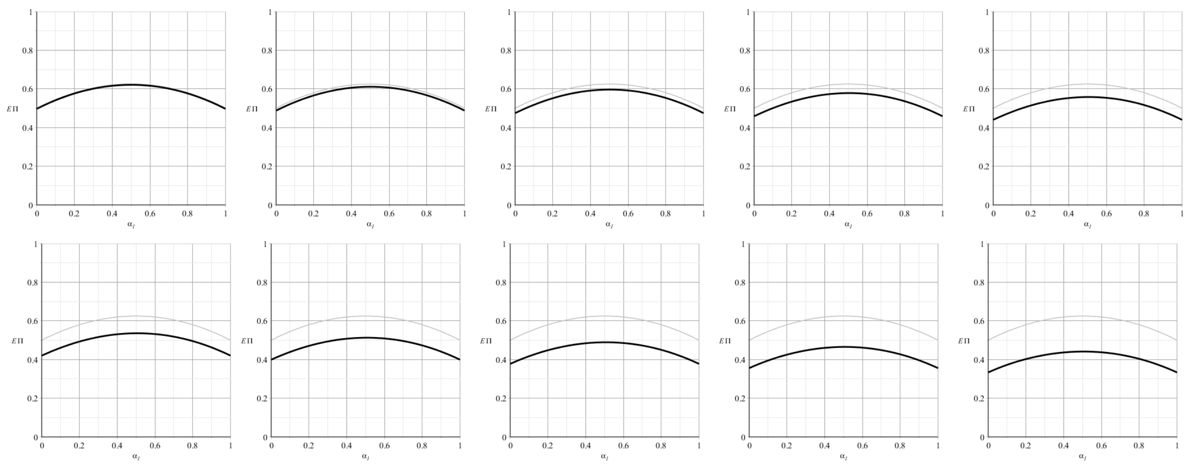


FIGURE 3.11: $E\big[\Pi^{H,U}(\alpha_i,\mu)\big]$ for $\mu=[0.1,0.2,0.3,...,1]$.

Under bilateral communication, as before we assume $\mu_1=\mu_2=\mu$ and we plot the three-dimensional functions of $E\big[\Pi^{H,B}(\alpha_1,\alpha_2,\mu_1,\mu_2)\big]$ for different values of μ. The graphs are presented in Figure 3.12. □

Proposition 3.4. *When agents have biased expectations, messages may have negative value.*

Proof. We first prove that, under unbiased expectations, messages always have positive value. We compare expected profits of dyads with different types of commmunication, given their structure.

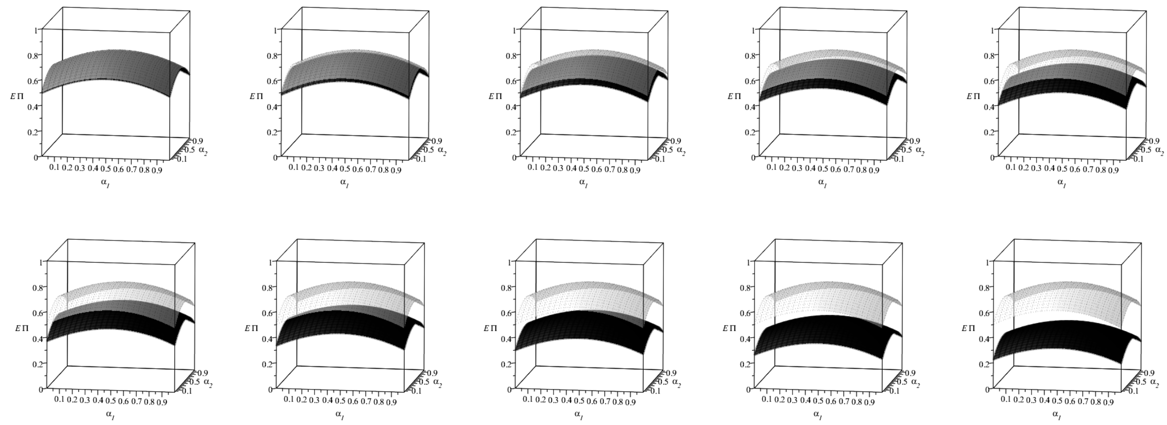


FIGURE 3.12: $E\big[\Pi^{H,B}(\alpha_1,\alpha_2,\mu)\big]$ for $\mu = [0.1, 0.2, 0.3, ..., 1]$.

In vertical dyads, unilateral communication is always better than no communication. In fact, $E\big[\Pi^{V,U}(\alpha_i)\big] > E\big[\Pi^{V,N}\big]$, i.e.

$$\frac{1}{2} + \alpha_i - \alpha_i^2 > \frac{1}{2}$$

is satisfied for each $\alpha_i \in (0,1)$.

In horizontal dyads, unilateral communication is always better than no communication. In fact, $E\big[\Pi^{H,U}(\alpha_i)\big] > E\big[\Pi^{H,N}\big]$, i.e.

$$\frac{1}{2}(1 + \alpha_i - \alpha_i^2) > \frac{1}{2}$$

is satisfied for each $\alpha_i \in (0,1)$.

Moreover, bilateral communication is always better than unilateral. In fact, $E\big[\Pi^{H,B}(\alpha_1,\alpha_2)\big] > E\big[\Pi^{H,U}(\alpha_1)\big]$:

$$\frac{1}{2}(1 + \alpha_1 - \alpha_1^2 + \alpha_2 - \alpha_2^2) > \frac{1}{2}(1 + \alpha_1 - \alpha_1^2)$$

for each $\alpha_1, \alpha_2 \in (0,1)$.

Now, we prove that the relation $E\big[\Pi^{V,U}(\alpha_i,\mu_i)\big] > E\big[\Pi^{V,N}(\mu_i)\big]$ is not satisfied for all $\mu_i \in (0,1)$. In Figure 3.13, we plot the following inequality

$$E\big[\Pi^{V,N}(\mu_i)\big] - E\big[\Pi^{V,U}(\alpha_i,\mu_i)\big] > 0$$

in the square $(0,1)^2$ as function of α_i and μ_i (the same graph is presented also in the main text).

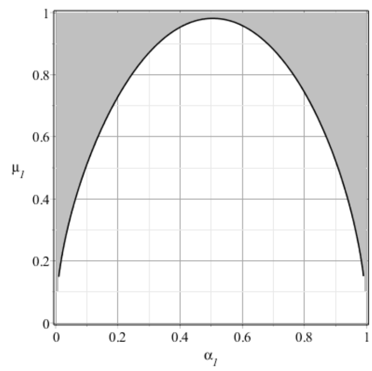


FIGURE 3.13: Area of (α_1, μ_1) under which communication has negative value for vertical dyads.

We now analyze horizontal structures. Again, without loss of generality we assume $\mu_1 = \mu_2 = \mu$. We first prove that $E\big[\Pi^{H,U}(\alpha_1, \mu_1, \mu_2)\big] > E\big[\Pi^{H,N}(\mu_1, \mu_2)\big]$ is not satisfied for all $\mu \in (0, 1)$. Hence, we plot condition

$$E\big[\Pi^{H,N}(\mu_1, \mu_2)\big] - \big[\Pi^{H,U}(\alpha_1, \mu_1, \mu_2)\big] > 0$$

in the Figure 3.14 (the graph has the same pattern of the previous one).

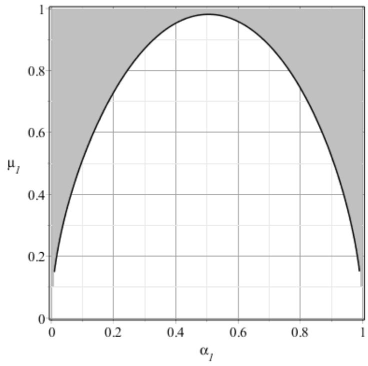


FIGURE 3.14: Area of (α_1, μ_1) under which communication has negative value for horizontal dyads.

Then, we prove also that $E\big[\Pi^{H,B}(\alpha_1, \alpha_2, \mu_1, \mu_2)\big] > E\big[\Pi^{H,U}(\alpha_1, \mu_1, \mu_2)\big]$ is not satisfied for all $\mu_i \in (0, 1)$. We assume $\mu_1 = \mu_2 = \mu$, and we plot the following inequality

$$E\big[\Pi^{H,U}(\alpha_1, \mu_1, \mu_2)\big] - E\big[\Pi^{H,B}(\alpha_1, \alpha_2, \mu_1, \mu_2)\big] > 0$$

as function of α_1 and α_2 for different values of μ proving that, as μ increases, unilateral communication tends to be more beneficial than bilateral. The result is presented in Figure 3.15.

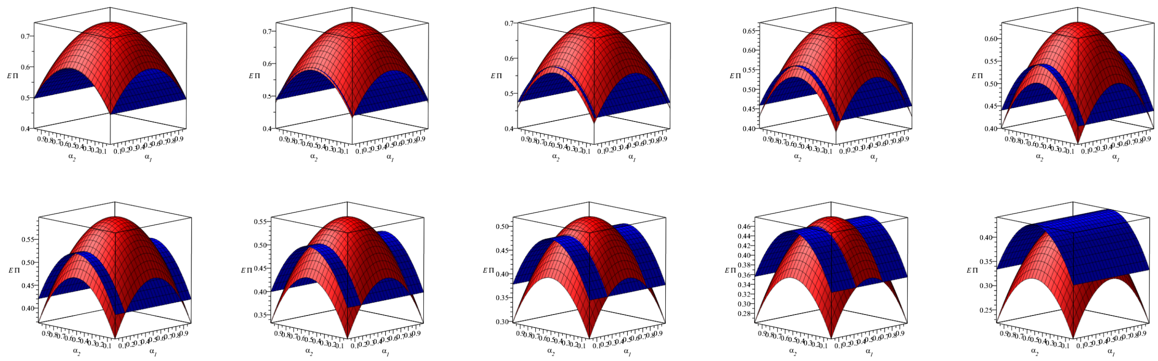


FIGURE 3.15: Area of (α_1, α_2), for $\mu = [0.1, 0.2, 0.3, ..., 1]$ under which unilateral communication (blue) is better than bilateral (red) for horizontal dyads.

Finally, we present a systematic overview of which communication type is the best one as function of α_1 and α_2 for $\mu_1 = \mu_2 = \mu$ and $\mu = [0.1, 0.2, 0.3, ..., 1]$. In green, we plot $E\big[\Pi^{H,N}(\mu_1, \mu_2)\big]$, in blue $E\big[\Pi^{H,U}(\alpha_1, \mu_1, \mu_2)\big]$, and in red $E\big[\Pi^{H,B}(\alpha_1, \alpha_2, \mu_1, \mu_2)\big]$. We see that, when the bias is high enough and α_1 and α_2 are peripheral enough, dyads are better off if team members do not exchange any message. The result is presented in Figure 3.16.

□

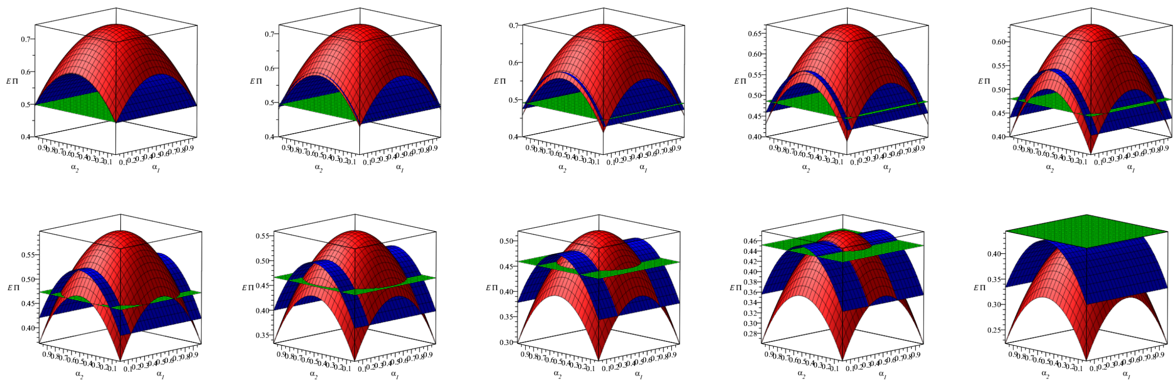


FIGURE 3.16: Comparison between different types of communication for horizontal dyads, for $\mu = [0.1, 0.2, 0.3, ..., 1]$: bilateral (red), unilateral (blue), and no communication (green)

Bibliography

[Alfaro et al., 2018] Alfaro, L., Bloom, N., Conconi, P., Fadinger, H., Legros, P., Newman, A.F., Sadun, R. and Van Reenen, J., 2018. "Come together: Firm boundaries and delegation", *NBER Working Paper No. 24603*.

[Bloom et al., 2012] Bloom, N., Sadun, R. and Van Reenen, J., 2012. "The organization of firms across countries", *The Quarterly Journal of Economics*, **127** (4), 1663-1705.

[Bolton et al., 2013] Bolton, P., Brunnermeier, M.K. and Veldkamp, L., 2013. "Leadership, coordination and corporate culture", *Review of Economic Studies*, **80** (2), 512-537.

[Brynjolfsson and Milgrom, 2013] Brynjolfsson, E. and Milgrom, P.R., 2003. "Complementarity in organizations", in Gibbons, R. and Roberts, J. (eds), *Handbook of Organizational Economics*, Princeton University Press, 680-731.

[Camerer and Lovallo, 1999] Camerer, C. and Lovallo, D., 1999. "Overconfidence and excess entry: An experimental approach", *American Economic Review*, **89** (1), 306–318.

[Everard, 2017] Everard, G., Jan. 20, 2017. "How to fail at Kickstarter - even if you get funded", *TechCrunch.com*, Web.

[Gary and Wood, 2010] Gary, M.S. and Wood, R.E., 2010. "Mental models, decision rules, and performance heterogeneity", *Strategic Management Journal*, **32** (6), 569-594.

[Gibbons and Henderson, 2013] Gibbons, R. and Henderson, R., 2013. "What do managers do?", in Gibbons, R. and Roberts, J. (eds), *Handbook of Organizational Economics*, Princeton University Press, 680-731.

[Gibbons and Roberts, 2012] Gibbons, R. and Roberts, J. eds., 2012. *The Handbook of Organizational Economics*, Princeton University Press.

[Heller and Winter, 2016] Heller, Y. and Winter, E., 2016. "Rule rationality", *International Economic Review*, **57** (3), 997–1026.

[Hong and Page, 2001] Hong, L. and Page, S.E., 2001. "Problem solving by heterogeneous agents", *Journal of Economic Theory*, **97** (1), 123–163.

[Hvide, 2002] Hvide, H.K., 2002. "Pragmatic beliefs and overconfidence", *Journal of Economic Behavior & Organization*, **48** (1), 15–28.

[Koellinger et al., 2007] Koellinger, P., Minniti, M., and Schade, C., 2007. "I think I can, I think I can: Overconfidence and entrepreneurial behavior", *Journal of Economic Psychology*, **28** (4), 502–527.

[Li et al., 1998] Li, J.C., Dunning, D., and Malpass, R.S., 1998. “Cross-racial identification among European-Americans: Basketball fandom and the contact hypothesis”, Unpublished manuscript, Cornell University, Ithaca, NY.

[LiCalzi and Surucu, 2012] LiCalzi, M. and Surucu, O., 2012. “The power of diversity over large solution spaces”, *Management Science*, **58** (7), 1408–1421.

[Mengel, 2012] Mengel, F., 2012. “On the evolution of coarse categories”, *Journal of Theoretical Biology*, **307**, 117–124.

[Prat, 1996] Prat, A., 1996. “Shared knowledge vs diversified knowledge in teams”, *Journal of the Japanes and International Economies*, **10** (2), 181-195.

[Rotemberg and Saloner, 1993] Rotemberg, J.J. and Saloner, G., 1993. “Leadership style and incentives”, *Management Science*, **39** (11), 1299-1318.

[Tanaka and Taylor, 1991] Tanaka, J.W. and Taylor, M., 1991. “Object categories and expertise: Is the basic level in the eye of the beholder?”, *Cognitive Psychology*, **23** (3), 457-482.

[Van Den Steen, 2005] Van Den Steen, E., 2005. “Organizational beliefs and managerial vision”, *The Journal of Law, Economics, & Organization*, **21** (1), 256-283.

[Wernerfelt, 2004] Wernerfelt, B., 2004. “Organizational languages”, *Journal of Economics & Management Strategy*, **13** (3), 461-472.