Università
Ca'Foscari
Venezia

Corso di Dottorato di ricerca
in Economia

Ciclo XXXI

Tesi di Ricerca

Three Essays on the Economics of Social Interactions

SSD: SECS-P/01

Coordinatore del Dottorato
ch. prof. Giacomo Pasini

Supervisore
ch. prof. Pietro Dino Enrico Dindo

Secondo Supervisore
ch. prof. Paolo Pin

Dottorando
Sebastiano Della Lena
Matricola 956245

Aknowledgments

After these three years of Ph.D. I feel to be a different person and I thank all the people I met for helping me on this journey.

I am deeply in debt to my supervisor Pietro Dindo. His guidance played a fundamental role in my development as researcher. Pietro was very patient in letting me investigate my research interests, to show me my weaknesses with kindness, and constantly help me to overcome them.

I deeply thank my second supervisor Paolo Pin, who in these years, with his example, made me passionate about research and inspired me. I am also grateful to Paolo for the invaluable help and guidance that he give me to the development of the third chapter of this thesis.

A special thank is due to Fabrizio Panebiaco, working with him allowed me to learn a lot and his company has been, and it is, very fulfilling both intellectually and personally. I have benefited from many conversations about research, and social interaction in general, with him.

I am very grateful to my committee members Sergio Currarini, Simon Weidenholzer, and Thierry Verdier, for the commitment and the effort that they have put in reading my work. Their comments helped me to significantly improve the three papers and encouraged me to keep on working with enthusiasm and passion.

I would like to thank Matteo Aggio, Caterina Pavese and Francesco Aldo Tucci most warmly for sharing the journey of these years together. Their friendship and support was fundamental for me during these years. I also thank all the Ph.D. students in Economics at Ca' Foscari University of Venice and Bocconi University, for the time spent together.

Finally, I thank my parents, Marcello and Marcella and my sister Giorgia, who in these years have always encouraged me, helped and, above all, endured me.

Contents

Introduction

This Doctoral thesis is divided in three chapters, each corresponding to a self-consistent theoretical paper in which, even if with different focus and methodology, the aim is to understand the economics beyond social interactions. In particular, I investigate how different strategic, informational or social environment affect the diffusion and evolution of agents' beliefs, preferences and norms.

The first chapter *"On the Interplay Between Norms and Strategic Environments"* is a joint work with Pietro Dindo and studies the role of different strategic environment for the dynamics of norms in a heterogeneous population divided into two cultural groups. The main contribution is to model norms as preferences over actions that modifies the payoff associated with playing according to it. We are able to reproduce different social outcomes, such as convergence toward the same social norm, the persistence of norms' heterogeneity, or even polarization of norms, while using the same norms formation model and depending on the type of strategic environment agents are exposed to during their adult life (e.g.*complements* vs *substitutes*)

In the second chapter *"Cultural Transmission with Incomplete Information: Parental Self-Efficacy and Group Misrepresentation"*, a joint work with Fabrizio Panebianco, we analyze, using the solution concept of self-confirming equilibrium, a cultural transmission model where parents have incomplete information about the the social structure and the efficacy of their vertical transmission efforts. Equilibria with wrong conjectures may arise. Main contributions are that in the short-run cultural complementarity instead of substitution holds and in the long-run the dynamics can display stable or unstable polymorphic equilibria, or just a stable homomorphic equilibrium.

The last chapter *"Non-Bayesian Social Learning and the Spread of Misinformation in Networks"* addresses the problem of the spread of permanent and temporary misinformation in a social network where agents interact to learn an underlying state of the world with a non-Bayesian social learning process. The main difference with the standard naive social learning is the continuous stream of new signals that agents receive at each period. Considering the permanent misinformation, pursuit by stubborn agents, we show that, despite receiving new signals every period, agents are not able to learn the underlying state of the word, nor to reach a consensus. The extent of deviation from the truth depends on a new measure of centrality "the updating centrality", that provides the key agents of the social learning process. Conversely, temporary misinformation, represented by shocks of rumors or fake news, has only short-run effects on the opinion dynamics. Our results are based on spectral graph theory techniques. In particular, using Perron-Frobenius theorem and Cheeger's inequality we show that the consensus among agents is not always a sign of successful learning. Moreover, the consensus time is increasing with respect to the "bottleneckedness" of the underlying network, while the learning time is decreasing with respect to agent's self-weights.

Chapter 1

On the Interplay Between Norms and Strategic Environments[1]

Abstract

This paper investigates the intergenerational dynamics of norms in a heterogeneous population divided into two cultural groups. In their adult life, agents are randomly matched to play symmetric 2 × 2 strategic games. Adults' norms, modeled as preference over actions, interact with material payoffs in determining best reply actions and thus Nash Equilibria. In turn, games influence norms because actions played in equilibrium reinforce the corresponding norm. At the end of their life, parents transmit their norms to offsprings who, in their youth, actively choose their own norm taking into account both the inherited norms and the norms of their peers. In our model, stable norms emerge as a steady state outcome of the joint dynamics of norms, actions, and socialization levels. We exploit this model to study the evolution of norms under different strategic environments: complements or substitutes. In general, complements and substitutes environment produce different social outcomes, namely full convergence and full divergence of norms, respectively. However, for specific choices of material payoffs and initial norms both partial and full convergence and divergence of norms can arise as stable outcome in both strategic environments.

Journal of Economic Literature Classification Numbers: C7, D9, I20, J15, Z1
Keywords: Evolution of Norms, Cultural Transmission, Endogenous preferences, Cultural Heterogeneity.

[1]This chapter is joint with Pietro Dindo, Ca' Foscari University of Venice, pietro.dindo@unive.it

From any given set of rules of conduct of the element will arise a steady structure (showing 'homeostatic' control) only in an environment in which there prevails a certain probability of encountering the sort of circumstances to which the rules of conduct are adapted. A change of environment may require, if the whole is to persist, a change in the order of the group and therefore in the rules of conduct of the individuals; and a spontaneous change of the rules of individual conduct and of the resulting order may enable the group to persist in circumstances which, without such change, would have led to its destruction.[Hayek 1967: 71]

1.1 Introduction

Some cultural traits, as language, result to be more homogeneous in the societies than others, like for example aggressiveness or effort choices at workplace. This is mainly due to the fact that incentives to coordinate are, often, stronger for some cultural traits than others. Nevertheless, although cultural assimilation is desirable, cultural heterogeneity is ubiquitous in different societies even when there are strong incentives to coordinate.[2]

In this paper, we consider the interplay between norms and Nash equilibrium outcomes of different strategic environments to explain the evolution of norms. We define norms as "mental representations of appropriate behavior" (Aarts and Dijksterhuis, 2003) or "internal standard of conduct" (Schwartz, 1977), namely they represent preferences over actions. By shaping agents' preferences and behaviors, different norms lead to different strategic outcomes (Akerlof, 1976; Young, 1998). At the same time, changes in the socio-economic environments can have an effect on the selection of norms (Hayek, 1967).[3]

Our approach focuses on three components: (i) the process of norms formation through socialization, (ii) the influence of norms on agents' preferences ordering, and thus on Nash equilibria, of one-stage games, (iii) the intergenerational transmission of norms and socialization levels as dependent on equilibrium actions played in one-stage games. The aim of the paper is to propose a model able to reproduce the emergence of different norms depending on the underlying strategic environment and to explore its properties for policy purposes.

[2]Bisin and Verdier (2012) offers a review of empirical examples of cultural heterogeneity and resilience of cultural traits. For example, the slow rate of immigrants' integration in Europe and US, the persistence of 'ethnic capital' in second- and third-generation immigrants or even cases of minorities' strongly attachment to languages and cultural traits.

[3]Our definition of norms differs from the literature of evolutionary game theory (**??**, for example) where a norm is broadly defined as an equilibrium of a strategic interaction.

Our analysis is aimed at reproducing different social outcomes, such as convergence toward the same social norm, the persistence of norms' heterogeneity, or even polarization of norms, while using the same norms formation model and depending on the type of strategic environment agents are exposed to during their adult life (e.g. *complements* vs *substitutes*). For example, we want to understand whether the material incentives to coordinate necessarily lead to a total assimilation of norms or, instead, they may leave space for multicultural society or, even, to the arising of an oppositional culture with the consequent separation of the minority. Relatedly, we investigate whether anti-coordinating environments are doomed not to be integrated. Moreover, we wonder if there exist socio-economic environments where material incentives make norms "disappear" in the long run. Understanding the relationship between the environment and the emergence of different cultures may help to better address policy issues. Modeling the interplay between actions, norms, and socialization levels enables us to show that different material payoffs, even within the same class of games, can lead to very different social outcomes. Policy intervention might benefit from this insight.

We consider a population divided into two communities, where individuals belonging to the same community are endowed with the same continuous cultural trait or personal norm. Individuals interact twice during their lives. First (while young), each agent forms a new norm taking into account both the inherited norm and inherited socialization level. The latter represents the willingness to conform to peers, as transmitted by parents. Then (in the adult age), agents are randomly matched to play symmetric 2×2 games. Payoffs, and thus best reply actions and equilibria, depend both on a material component and on an immaterial one, the latter being norm dependent. Thus, norms parametrize agent preferences over material payoffs. At the end of their lives (old age), each agent transmits a norm and a socialization parameter to his offspring. The transmission is moved by *cognitive dissonance* and *cultural substitution*. *Cognitive dissonance* is the tendency of agents to have consistency between behavior (action) and norms (preferences over actions) and it is a key assumption to model the feedback from the strategic environment to the norms' transmition. *Cultural substitution* captures the idea that the vertical socialization level of offsprings negatively depends to the diffusion of parents' behavior in the population. The assumption of *cultural substitution* allows us to study the evolution of socialization's levels but do not affect results about norms dynamics.

In Section 1.2.1, we analyze the norm formation mechanism of the youth age. At this stage, we study how, interacting with peers, agents of both communities symmetrically form the norms that will affect the payoff structure of games played in the adult age. The main intuition is that children are not passive during the transmission process; on the contrary, they are responsible for the formation of their own norms. We follow Kuran and Sandholm (2008) and assume that in choosing the norm, each agent faces a trade-off between his observable[4] inherited norm and a social coordination payoff, minimizing a loss function. A socialization parameter describes the strength of such trade-off. With respect to the previous literature, we introduce heterogeneity in the socialization parameter, which determines the weight put on the inherited norm (vertical socialization) and the weight put on the entire population average norm (horizontal socialization).[5]

In Section 1.2.2, we model the effect of norms on different strategic environments. We interpret each agent's norm as the preference for a particular behavior that modifies the payoff associated with playing according to it. The intuition is that some behavior (action) can be more or less in line with the personal norm of one agent. Each strategic environment is a symmetric 2×2 game, which is meant to be representative of tasks that people can face in their adult age. We mainly discuss two different kind of norms associated with different strategic environment. In particular, we have in mind a linguistic dilemma as an example of strategic environment with *complements* (i.e. the material incentive is to coordinate to the same language) and the choice of being aggressive or not in a competition for a shared resource for what concern strategic environment with *substitution* (i.e. the material incentive is to ant-coordinate on the level of aggressiveness, as in the classical hawk-dove game).[6] In their adult age agents can face several games, of the same type, described by a distribution of payoffs. Agents of each community strategically interact with agents belonging to both groups, in random matching games. We propose a multiplicative interaction between norms and material payoffs. If the norms are neutral the payoffs of the game are equal to material payoffs and agents play the original 2×2 game. If norms assume extreme values agents stick to the associated action, giving no importance to material payoffs. When norms have intermediate values, there is a trade-off between the material consequence of actions and playing according to the behavior associated with such norms. Relatedly, we derive the possible Nash equilibria as depending on the tension between material payoffs and norms over behavior.

In Section 1.2.3, in order to characterize the feedback between the strategic environment and norms, we study the transmission of norms from old to young and the evolution of the socialization parameter. We exploit the tendency of people to seek consistency between preferences over actions and behavior. In particular, we let agents move their norm toward the action predominantly played in the adult age, to reduce *cognitive dissonance*, as in Kuran and

[4] We assume complete information.

[5] We refer to Cavalli-Sforza and Feldman (1981); Bisin and Verdier (2001) for the terminology.

[6] We also discuss how, depending on the circumstances, some norms can be associated to both *complements* or *substitution* environments, for example the optimal effort level in team work strongly depends on the nature of the task and if there is the possibility to free-ride on the team-mate's effort.

Sandholm (2008). Agents transmit this modified norm to their offspring. The choice of socialization also depends on actions played in the adult age and is moved by *cultural substitution* (Bisin and Verdier, 2001). Namely, we assume that parents have fewer incentives to vertically socialize their children when their behavior is widely spread in the population.

In Section 1.3, we derive all possible long-run outcomes. We consider two limit cases: uniform and point distribution of one-stage game payoffs. On the one hand, with uniform distribution of payoffs environments with *complements* or *substitutes* produce very different social outcomes. In a social environment with *complements*, cultural assimilation, i.e. both communities share the same norm and behavior, emerges as stable steady state. On the other hand, *substitution* in the material incentives leads the emergence of oppositional cultures, the strong polarization of norms and behavior of agents belonging to different communities. The different steady states also have different socialization levels. Under assimilation, agents have a maximum horizontal socialization level, and thus a minimum vertical socialization. Under polarization, the horizontal socialization level is close to its minimum, and the vertical one is close to its maximum. Interestingly, when there is polarization of norms, the larger the majority, the farther away are both norms and socialization levels. Indeed, in order to stick to its preferred behavior (different from the one of the majority) the smaller the minority is, the higher the vertical socialization becomes.

Results change when the one-stage game payoffs is not uniform. Provided agents play always the same game (singular payoff distribution) and depending on the initial norms, in both *complements* and *substitutes* environments it is possible to converge toward cultural assimilation or to diverge and having the arising of oppositional cultures or separation. Moreover, we show that even partial convergence or partial polarization can be sustained in games with *complements* and *substitutes*, respectively. In these cases, one community has a norm so strong as to generate a dominant strategy while the other does not have such a strong norm and best replies to the dominant action only by looking at material payoffs.

Section 1.4 discusses about general payoff distributions, the role of the assortativity on the matching process for the steady state outcome and speed of convergence, and discuss possible further development of the model allowing for mixed (*complements* and *substitutes*) environments.

Section 1.5 concludes the paper and discusses possible extensions.

1.1.1 Literature Review

In recent years a very wide literature about norms and their effect in socio-economic outcomes has emerged. Many works focus on the relationship between norms (or culture) and coordination. Acemoglu and Jackson (2014) study the evolution of a cooperation norm; Dalmazzo et al. (2014) present conditions under which harmful cultural traits can persist in a community; Michaeli and Spiro (2017) address the arising of biased norm when agents, with pressure to conform to each other, play coordination game; Carvalho (2016) shows how cultural constraints can lead to miscoordination. Along this research line, we should consider Tabellini (2008), where agents, matched together to play a Prisoner Dilemma, face a trade-off between individual values (inherited from parents) and material incentives. The main contributions of our work with respect to these papers are that we study the outcome for different classes of games at once, and that the norm formation process depends both on the imitation of peers (horizontal socialization) and parents' transmission (vertical socialization).

The literature about cultural transmission was initiated by Cavalli-Sforza and Feldman (1981) and, in economics, by Bisin and Verdier (2001), where the evolution of cultural traits is the result of parent's socialization choices. Socialization can be vertical (parents), horizontal (peers), and oblique (role models). Along these lines, Bisin and Verdier (2017) study the joint evolution of culture and institutions. In our paper, the socialization is vertical, when parents transmit their preferences to offsprings, and horizontal, when peers interact together to form new norms. In our model the transmitted cultural traits are continuous, as in Panebianco (2014). For a complete theoretical and empirical survey on cultural transmission literature see Bisin and Verdier (2012).

For what concern the effect of norms on the payoff structure, this paper refers to a specific behavioral literature (López-Pérez, 2008; Kessler and Leider, 2012; Kimbrough and Vostroknutov, 2016) where actions are moved by the will to adhere to a norm.[7] The main difference is that in our paper agents are affected by a group-specific norm, not necessarily equal for the whole society, so that different players can be subject to different norms.

The concept of *cognitive dissonance* that we use for the dynamics of preferences was introduced in economics by Akerlof and Dickens (1982). Kuran and Sandholm (2008) and Calabuig et al. (2014), whose norm formation and norm dynamics are close to ours, have also elements of cognitive dissonance in the updating of norms. In particular, our contribution can be seen

[7] An alternative viewpoint is that norms imply preferences for a certain distribution of outcomes, i.e. uniform across players.

as an extension of Kuran and Sandholm (2008) where agents, endowed with their norms, interact in strategic environments, and where the dynamics of norms depends on the interaction between norms and the related equilibrium outcome of games. If we switch-off feedbacks of equilibrium action on transmitted norms and horizontal socialization, our model boils down to a discrete time version of Kuran and Sandholm (2008). Having these feedbacks changes the results drastically, for example we are able to reproduce cultural heterogeneity even with fixed communities and complete interaction.

Concerning the applications, our model talks to the literature of identities and oppositional cultures pioneered in economics by Akerlof and Kranton (2000); for example Kuran and Sandholm (2008) study the tension between cultural integration and multiculturalism, Bisin et al. (2011) focus on the reason that leads to the presence of oppositional cultures, Olcina et al. (2017) address the problem of minorities embedded in a relationships network who decide whether or not to be assimilated to the majority norm. Our main contribution with respect this literature is to make explicit the effect of different strategic environments.

Our results can be applied even in the framework of learning and opinion dynamics (DeGroot, 1974; DeMarzo et al., 2003; Golub and Jackson, 2010, 2012) where there is the tension between reaching consensus and disagreement. For example, Yildiz et al. (2013) find in the presence of stubborn agents the reasons of disagreement, Golub and Jackson (2010) study the general conditions for reaching the consensus in a network, and **?** show how an endogenous network structure can lead to opinion's polarization. In this framework, according to our model, agents form their opinion taking into account both their previous opinion and the one of others. Then, when they are supposed to take decisions, they are affected by both opinions and material rewards. Thus, they update and transmit new opinions taking into account also the experience gained through interaction. The main insight of our work with respect to this literature is that the interplay between material incentives and opinions may be crucial for leading to a consensus or to disagreement.

1.2 The Model

In this section, we model the interplay between norms and strategic environments, as represented by games. Consider a non-atomic society as a continuum of individuals of mass 1. Agents in the society are divided into two communities $\mathcal{I} = \{1, 2\}$. Define $\eta \in \left(\frac{1}{2}, 1\right)$ the size of the majority and assume, without loss of generality, that community 1 is the majority. Agents belonging to the same community are assumed to be equal and $i \in \{1, 2\}$ is the representative agent of each community.

Each period time $t \in \mathbb{N} \cup \{0\}$ represents a generation of agents. We divide the life into three different sub-periods. In *Stage (i)*, youth, the social coordination game that microfunds the choice of personal norms takes place, in *Stage (ii)*, adult age, agents interact by playing games whose payoffs are determined also by personal norms, in *Stage (iii)*, old age, norms are transmitted to the next generation.

Stage (i) When young, members of the two communities are endowed with type-specific observable personal norms $\theta_t = (\theta_{1,t}, \theta_{2,t}) \in [0,1]^2$ and flexibility parameters $f_t = (f_{1,t}, f_{2,t}) \in [0,1]^2$, both characteristics are inherited by the previous generation. Interacting together, young symmetrically choose *ex-post* personal norms $x_t = (x_{1,t}, x_{2,t}) \in [0,1]^2$ taking into account both inherited norms and a preference for conformity.

Stage (ii) During their adult age, agents interact in a strategic environment. Agents are randomly matched in pairs to play several symmetric 2×2 games. Different games are available in the same period and each game is played with probability γ. Each agent plays with members of both communities, namely η time against the majority and $1-\eta$ against the minority. Games and population match are drawn from independent distributions. Norms influence total payoffs and the Nash equilibrium actions emerge as the response of both material payoffs and personal norms x_t. $E_{\eta,\gamma}[A_t] = (E_{\eta,\gamma}[A_{1,t}], E_{\eta,\gamma}[A_{2,t}]) \in [0,1]^2$ is the vector of average actions in period t, where $E_\rho[.]$ is the expectation operator with respect to measure ρ.

Stage (iii) At the end of their life, every agent reproduces asexually, giving birth to one child. At this stage, parents transmit new norms and choose how much to socialize their offsprings. During the transmission, parents are assumed to be myopic and to be not able to anticipate the future utility of the offspring, therefore the transmission is fully mechanical. We model the feedback from the environment (game) to norms as a cultural transmission where the Nash equilibrium action most played in the game leads to the inherited personal norms of the new generation. In particular, we consider two forces: *cognitive dissonance*, for the choice of the norm to be transmitted, and *cultural substitution*, for the choice of the socialization parameter f.[8]

We can think about this model as a three-period overlapping generation model (young, adults and old).

[8]We provide definitions and details about *cognitive dissonance* and *cultural substitution* in Section 1.2.3.

We start our analysis, in the next section, from the illustration of the formation of norms in the young age. Next, we consider the effect of norms on the payoffs of games played in agent's adult lie. Finally, we characterize the transmission process. In the first two sections, we avoid the time index to simplify the notation.

1.2.1 Young Age Norms' Formation

In this section, we model agents choice of *ex-post* norms $x \in [0,1]^2$, stemming from the inherited norms θ and horizontal socializations f. In our model, young agents (children) are active in choosing their own personal norms.

As in many cultural evolution paper (e.g. Kuran and Sandholm, 2008) agent *ex-post* norms x are the result of social interaction. The general idea is that each agent choice of a norm is affected both by his inherited norm and by the average *ex-post* norm chosen by his peers.[9] Since we have two communities θ, f and x are two dimensional vectors. In particular, we assume that the utility of a generic agent of one community $i \in \{1,2\}$ is:[10]

$$u_i(x, \theta_i, f_i) = -\underbrace{f_i(x_i - E_\eta[x])^2}_{social\ coordination} - \underbrace{(1-f_i)(x_i - \theta_i)^2}_{group\ (or\ family)\ identity}$$

where $E_\eta[x]$ is taken over the distribution of individual characteristics and it is the average chosen norm.[11] The parameter f_i, which can differ across communities, is the horizontal socialization. $(1 - f_i)$ is thus the vertical socialization. The utility function captures the tension between inherited preferences and coordinating with others. When choosing a personal norm agents want to pick one not too different from the one of their peers, depending on their horizontal socialization parameter.[12]

Given the distribution of θ and f across the population, the *ex-post* personal norm is defined by the unique symmetric Nash equilibrium of the social interaction game, where agents of the same type choose the same ex-post personal norm. The *ex-post* norm of an agent belonging

[9] This is consistent with sociological literature about social norms, see **?** for a survey.

[10] Notice that, in principle, agents in the same community can make different choices. However, we focus only on symmetric choices for all the agents of the same community and thus, with an abuse of notation, we use only the community index from the beginning.

[11] $E_\eta[x]$ can be seen as descriptive norm (**?**). Notice that since agents are myopic in their youth they are not able to anticipate future payoffs and thus form their norms taking in consideration only parents and peers pressure and not the subsequent strategic environment.

[12] This formulation is exactly equivalent to the *conformity game* played by children in Vaughan (2013). Moreover, beauty contest like utility function, introduced in economics by Morris and Shin (2002), is widely used both in the literature of evolution of cultural traits (Kuran and Sandholm, 2008, among others) as well as in network economics for opinion or belief learning and dynamics (Golub and Jackson, 2012; **?**), where it can be seen as a micro-foundation of the so called De Groot model (DeGroot, 1974)

to community i depends on both the distribution of inherited norms θ and on the distribution of socialization parameters f.

Proposition 1

For all θ and f, there exists a unique symmetric Nash Equilibrium of the norm formation game with

$$x_i = f_i \left(E_\eta \left[\theta\right] - \frac{cov_\eta[f,\theta]}{(1 - E_\eta[f])} \right) + (1 - f_i)\theta_i \quad \text{for } i = 1, 2 \tag{1.1}$$

The average norm is

$$E_\eta[x] = E_\eta[\theta] - \left(\frac{E_\eta[f]}{1 - E_\eta[f]} \right) cov_\eta[f, \theta] \tag{1.2}$$

Proof. In the Appendix. □

The result is a generalization of Kuran and Sandholm (2008). In that setting, f is the same for both groups, $cov[f,\theta] = 0$, and thus (1.2) $E_\eta[x] = E_\eta[\theta]$: the population average *ex-post* norm is equivalent to the population average inherited norm. In our model, the heterogeneity of horizontal socialization introduces a distortion in the distribution of *ex-post* norms. If the covariance between socialization parameters and norms is zero, then the average *ex-post* norm in the society is exactly the average of inherited norms. On the other hand with positive or negative covariance, even a minority, if enough rigid, can make her group norms prevail.
Since we have only two types of communities[13], we can represent the equilibrium norms as a convex combination of inherited norms as follows.

Corollary 1.1

The Nash Equilibrium of Proposition 1 can be written as

$$\begin{cases} x_1 & = p_1\theta_1 + (1 - p_1)\theta_2 \\ x_2 & = p_2\theta_1 + (1 - p_2)\theta_2 \end{cases}, \tag{1.3}$$

where $p_1 = \frac{(1-f_1)(1-f_2(1-\eta))}{1-f_1\eta-f_2(1-\eta)} \in (0,1)$, $p_2 = \frac{f_2\eta(1-f_1)}{1-f_1\eta-f_2(1-\eta)} \in (0,1)$ *and* $p_1 > p_2$ *for all* η, f_1, f_2.

Each agent, as result of the social interaction, chooses a norm that is a convex combination between her initial norm and the one of the other community. Weights depend on both types socialization parameters and the majority size η. By taking the difference of p_1 and p_2, it can

[13]However that Proposition 1 is valid for any number of communities.

be easily seen that p_1 is always greater that p_2. Thus if $\theta_1 > \theta_2$, then $x_1 > x_2$ (and viceversa): it is not possible to have a switch of ordering between *ex-ante* and *ex-post* norms.

1.2.2 Nash Equilibria for Normal Form Games with Norms

In this section, we model how norms change the payoffs of each one-stage game and study the implication on the game's (pure) Nash equilibria. Agents in the adult age are called to make strategic decisions. Their choice is affected both by material and immaterial payoffs. The latter are represented by the willingness to choose an action as indicated by their norms. Some norms are often associated with cooperative environment while others with competitive ones. Before to proceed with the analysis we provide anecdotic examples about norms associated with environments with *complements*, *substitutes*, or both.

Language (complements) When people interact in a multicultural environment, they have to choose the language to use. On the one hand, there are evident "material" incentives to coordinate on the same language. On the other hand, agents can have different preferences in using a specific language (norms), depending on agents ability to speak the two languages or other idiosyncrasies. The interaction of such preferences with the incentive for coordination may result in different Nash equilibria. If the norm is mild, one possibility is to use the native language when meeting a person of the same community and the leading language otherwise. If the norm is "strong enough" the game could instead become an anti-coordination one, when each member of a community uses only its most preferred language.

Aggressiveness (substitutes) In competitive interactions for shared resources the material incentives is to anti-coordinate, the optimal action is to be aggressive when the other agent is not, and viceversa (see hawks-doves type of games). In this case, the underlying material environment is an anti-coordination game, but norms can transform it in a coordination one. Namely, very pacific types would always behave peacefully independently on incentives, and viceversa.

Work Ethics (mixed) In interacting at the workplace people may face both an environment with *complements* and with *substitutes*. If we consider a work task that needs a team effort to be accomplished and there is no reward if both agents do not exert a high level of effort, agents have incentives to coordinate and the game is with *complements*. On the contrary, easy tasks that can be accomplished with the effort of only one agent open the doors for freeriding and the game is with *substitutes*. This two examples can be though as, on the one hand, a tough economy where resources are extractable at high labor cost in which agents have to cooperate to survive (*complements*), and, on the other hand, a flourishing economy where there are abundant and easily extractable resources, in which some agents have the chance to freeride (*substitutes*).

In the general case, agents play different games in each round. In this section we shall characterize how norms change payoffs and thus the Nash Equilibria of each of these games. We represent the tasks that individuals face with symmetric bimatrix games where norms x interact with material payoffs and lead to the total, material plus immaterial, payoff. For each bimatrix game Γ, the set of players is $\mathcal{N} = \{r, c\}$ and the action space is defined as $\mathcal{A} = \mathcal{A}_r \times \mathcal{A}_c$ where $\mathcal{A}_r = \mathcal{A}_c = \{1, 0\}$ is the set of actions available for each player (e.g. language A or language B, being aggressive or not, high effort or low effort at the workplace). Each player belongs to one of the two communities and i_r (j_c) is the community of player r (c). The material payoff function of the symmetric bimatrix game is $\pi : \mathcal{A} \to \mathbb{R}^2$, where: $\pi_r(1,1) = \pi_c(1,1) = a$, $\pi_r(0,0) = \pi_c(0,0) = d$, $\pi_r(1,0) = \pi_c(1,0) = c$ and $\pi_r(0,1) = \pi_c(0,1) = b$.

Norms chosen in previous period x belong to the closed interval $[0,1]$, while the actions of the game are binary and are associated with extreme norm values 0 or 1.[14] We consider norms as preferences over behavior, namely, the more the behavior deviates from the norms, the more the agents face a loss of utility.[15] Therefore, the real perceived payoffs are, in general, different from just the material payoffs. In particular, for each $A = (A_r, A_c) \in \mathcal{A}$ player r has the following payoff function with norms

$$\Pi_r(A; x_{i_r}) = [A_r x_{i_r} + (1 - A_r)(1 - x_{i_r})]\pi_r(A_r, A_c). \tag{1.4}$$

The same modification of the material payoff holds also for player c. Equation (1.4) implies that the payoff of a game with norm x_{i_r} is $\Pi_r(1, A_c; x_{i_r}) = x_{i_r}\pi_r(1, A_c)$ when $A_r = 1$ and $\Pi_r(0, A_c; x_{i_r}) = (1 - x_{i_r})\pi_r(0, A_c)$ when $A_r = 0$. Table 1.1 represents the payoff matrix associated with Γ.
When $x > 1/2$ the norm is in favor of action 1 so that agent total payoff for playing action 1 is larger than when the norm is $x = 1/2$ and than when the norm is in favor of action 2, $x < 1/2$. The larger the norm the larger such an influence. If a norm is extreme ($x = 0$ or $x = 1$), the agent always plays the associated action, thus giving no importance to material

[14] The action can be though depending on the situation as use language A or language B, use min effort or max effort, etc.

[15] Notice that we could express this ordering of preferences over consequences in the framework of psychological games (**??**, among others) where players have belief-dependent motivations (such as intentions-based reciprocity, emotions, or concern with others' opinion) in our case agents would have *ex-ante* beliefs about their own action (norms) and care about the consistency of these beliefs and the chosen actions.

		Agent c	
		1	0
Agent r	1	$x_{i_r}a, x_{j_c}a$	$x_{i_r}c, (1-x_{j_c})b$
	0	$(1-x_{i_r})b, x_{i_m}c$	$(1-x_{i_r})d, (1-x_{j_c})d$

Table 1.1: Bimatrix game with norms as preferences over behavior

payoffs. In these cases we say that the norm is *strong*. When instead $x = \frac{1}{2}$ the transformation has no effect on best replies and the norm is said to be *neutral*. For example, if r is neutral, then $\Pi\left(A; \frac{1}{2}\right) = \frac{1}{2}\pi(A)$ independently on played action and thus with no effect on best replies.

Modeling the effect of norms as in (1.4) we capture the idea that norms change preferences over material payoffs, as in Bisin et al. (2004) and Tabellini (2008) but, at the same time, when norms are extreme (0 or 1), the transformation is consistent with the interpretation of Carvalho (2016) where norms restrict agents' strategy set.[16]

Since we are interested in studying the evolution of norms when agents are exposed to different strategic environments, in particular when in absence of norms actions are *complements* or *substitutes*, we restrict the ordering of the material payoffs a, b, c, d. Defining $\bar{b} = \frac{b}{a+b}$ the material force that leads out from the equilibrium $(1,1)$ and $\bar{d} = \frac{d}{c+d}$ the force that pushes toward the equilibrium $(0,0)$, it is possible to categorize the possible games with material payoffs $\pi(A)$ as $\Gamma(\bar{b}, \bar{d})$. In particular

1. Coordination (*Complements*): $\bar{b} < \frac{1}{2} < \bar{d}$
2. Anti-Coordination (*Substitutes*): $\bar{d} < \frac{1}{2} < \bar{b}$.[17]

Therefore, we can define the game with norms as $\Gamma(\bar{b}, \bar{d}, x)$. When $\bar{b} < \frac{1}{2} < \bar{d}$, we say that the society interact in a environment with *complements*. When $\bar{d} < \frac{1}{2} < \bar{b}$, agents face an environments with *substitutes*. Notice that, similarly to norms, also $\bar{b}$ and $\bar{d}$ belongs to the interval $[0,1]$. However, they represents material, instead of moral, incentives. If $\bar{b} = \bar{d} = \frac{1}{2}$, then there are no material incentives.

[16] An agent that takes into account both material and moral payoffs as in (1.4) can be seen as "Homo Moralis" in the language of Alger and Weibull (2013). Moreover, the functional form for payoffs (1.4) is also consistent with one commonly used in the behavioral literature on social norms (López-Pérez, 2008; Kessler and Leider, 2012; Kimbrough and Vostroknutov, 2016) where a cost function c of violating the norm is subtracted to the material payoff: $\Pi_r(A_r, A_c; x_{i_r}) = \pi_r(A_r, A_c) - c(x, A, \pi)$. Indeed, with $c(x, A, \pi) = \pi_r(A_r, A_c)(A_r + x_{i_r}(1 - 2A_r))$ we get exactly equation (1.4).

[17] Notice that restricting the ordering of $\bar{b}, \bar{d}, \frac{1}{2}$ can be used to characterize even Prisoner Dilemma and Efficient Dominant Strategy Equilibrium games, where $\frac{1}{2} < \min\{\bar{b}, \bar{d}\}$ and $\frac{1}{2} > \max\{\bar{b}, \bar{d}\}$, respectively. We do not consider them because in this work we want to focus on *complements* vs *substitutes*.

We can now proceed with the equilibrium analysis. For simplicity, we assume complete information about material payoffs, norms, and rationality of agents and thus we can use (pure actions) Nash equilibria. The equilibrium analysis relies on the double effect on moral and material incentives. Moral incentives depend on the consistency between norms and actions. Therefore, the final decision depends on the strength of the norm as compared to the two threshold $\bar{b}$ and $\bar{d}$. $\bar{b}$ establishes the minimum strength of the norm for action 1 to be played when the opponent plays 1. $\bar{d}$ establishes the maximum strength of the norm for action 0 to be played when the opponent plays 0.[18]

Define $\hat{A}(\bar{b}, \bar{d}, x) := \hat{A}_r(\bar{b}, \bar{d}, x_{i_r}) \times \hat{A}_c(\bar{b}, \bar{d}, x_{j_c})$ the set of Nash Equilibria with norms.

Proposition 2 *In a Bimatrix Game with norms* $\Gamma(\bar{b}, \bar{d}, x)$*:*

- *If* $x_{i_r}, x_{j_c} > \bar{b}$*, then* $(1,1) \in \hat{A}(\bar{b}, \bar{d}, x)$*.*
- *If* $x_{i_r}, x_{j_c} < \bar{d}$*, then* $(0,0) \in \hat{A}(\bar{b}, \bar{d}, x)$*.*
- *If* $x_{i_r} > \bar{d}$ *and* $x_{j_c} < \bar{b}$*, then* $(1,0) \in \hat{A}(\bar{b}, \bar{d}, x)$*.*
- *if* $x_{i_r} < \bar{b}$ *and* $x_{j_c} > \bar{d}$*, then* $(0,1) \in \hat{A}(\bar{b}, \bar{d}, x)$*.*

Proof. We solve the generic game with norms as described in Table 1.1. The best-replies are:
$\hat{A}_r(A_c = 1; \bar{b}, \bar{d}, x_{i_r}) = \begin{cases} 1 \; if \; x_{i_r} > \bar{b} \\ 0 \; if \; x_{j_c} < \bar{b} \end{cases}$ and $\hat{A}_r(A_c = 0; \bar{b}, \bar{d}, x_{i_r}) = \begin{cases} 1 \; if \; x_{i_r} > \bar{d} \\ 0 \; if \; x_{j_c} < \bar{d} \end{cases}$.

Since the game is symmetric we do not compute the best reply for agent j. Looking for the fixed-point of the best replies we find the Nash Equilibria. □

Figures 3.2 represents the actions played in equilibrium by two agents belonging to different communities as a function of their norms and for different strategic environments: *complements* or *substitutes*. Nash equilibria depend on the position of threshold values $\bar{d}$ and $\bar{b}$. Figure 3.2(a) represents a game with *complements*, $\bar{b} < \frac{1}{2} < \bar{d}$. Figure 3.2(b) represents a game with *substitutes*, $\bar{d} < \frac{1}{2} < \bar{b}$. The main diagonal represent the action played by agents of the same community, since the norm is the same. When the action is marked with the subscript $*$, it is a dominant strategy. As expected, for x_{i_r}, x_{j_c} in the neighborhood of neutral norms ($x_{i_r} = x_{j_c} = \frac{1}{2}$), the games have the same equilibria as the corresponding game without norms,

[18] This is consistent, even if in a totally different framework, with Eshel et al. (1998), who found that the imitation dynamics depends only upon the values α and β which are strictly related respectively with our $\bar{b}$ and $\bar{d}$.

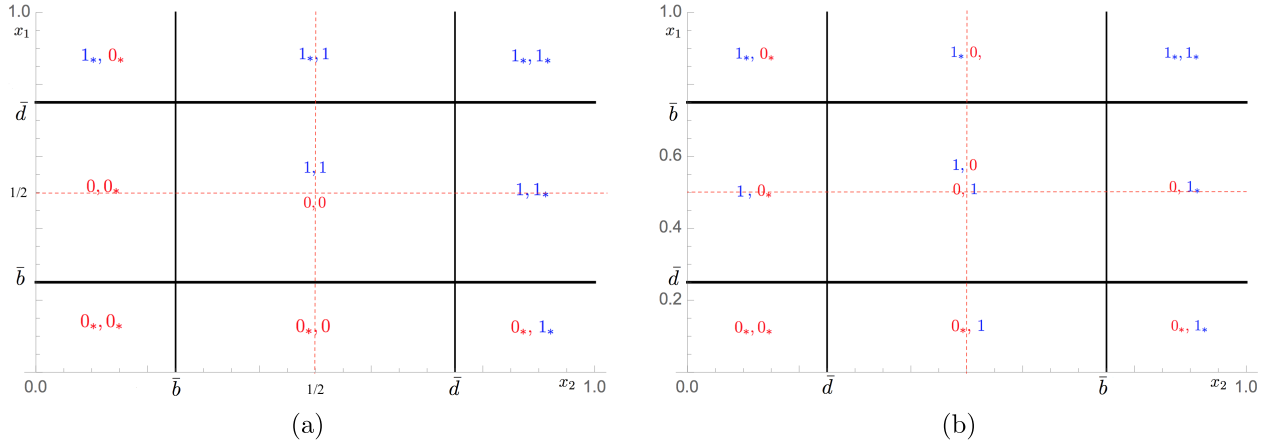


Figure 1.1: Equilibria Space of Games with (a) *Complements* $\bar{b} < \frac{1}{2} < \bar{d}$ and (b) *Substitutes* $\bar{d} < \frac{1}{2} < \bar{b}$.

while as x_{i_r}, x_{j_c} move away from $\frac{1}{2}$ the games have have different equilibria.

Notice that when $x_{i_r} > \max\{\bar{d}, \bar{b}\}$ or $x_{i_r} < \min\{\bar{d}, \bar{b}\}$, r has a dominant strategy, respectively playing 1 or playing 0. Otherwise, if $\min\{\bar{d}, \bar{b}\} < x_{i_r} < \max\{\bar{d}, \bar{b}\}$, then r has not a dominant strategy and reacts to the action of c. Thus if both x_{i_r} and x_{j_c} are between $\min\{\bar{d}, \bar{b}\}$ and $\max\{\bar{d}, \bar{b}\}$, then we have multiple equilibria.

For all games, we can distinguish different areas in the space of personal norms of the two players.[19] In particular, there are five main areas to keep in considerations. For extreme norm values, the games have a unique equilibrium: in $R_{1_*,1_*}$, $R_{1_*,1}$ and $R_{1,1_*}$, the x's approach 1 and the equilibrium is $(1, 1)$; in $R_{0_*,0_*}$,$R_{0_*,0}$ and $R_{0,0_*}$ the xs tend to 0 and the equilibrium is $(0, 0)$. Only in 2 of these 6 regions both players play a dominant strategy. In $R_{1_*,0_*}$ and $R_{0_*,1_*}$, norms are polarized and Nash equilibria are, respectively, $(1, 0)$ and $(0, 1)$ for all type of strategic environments. With intermediates value of x, in R, the game has multiple equilibria. In particular, if $\bar{b} < \bar{d}$ we have a coordination game with the two equilibria $(1, 1)$ and $(0, 0)$; if $\bar{b} > \bar{d}$ we have an anti-coordination game and the two equilibria are $(1, 0)$ and $(0, 1)$.

We shall assume that in each period agents play different games. In particular, we name γ a probability density on the space of vectors $(\bar{b}, \bar{d})$ and assume that the game with payoffs $\Gamma(\bar{b}, \bar{d}; x)$ is played with probability $\gamma(\bar{b}, \bar{d})$. In order to study the effect of different strategic environment on norms dynamics, we further assume that games played belong always to the same environment. Namely, in complements environments $\bar{b} \in \left[0, \frac{1}{2}\right]$ and $\bar{d} \in \left[\frac{1}{2}, 1\right]$, viceversa

[19]See in the appendix Corollary 2.1, for a formal definition of each region.

for substitutes environments $\bar{b} \in \left[\frac{1}{2}, 1\right]$ and $\bar{d} \in \left[0, \frac{1}{2}\right]$. In particular, we focus our analysis on the two extrema cases: (i) γ is a uniform distribution on sets of $(\bar{b}, \bar{d})$ with support $\left[0, \frac{1}{2}\right] \times \left[\frac{1}{2}, 1\right]$ for complements and $\left[\frac{1}{2}, 1\right] \times \left[0, \frac{1}{2}\right]$ for substitutes. (ii) γ is a point distribution of an element of the set $\left[0, \frac{1}{2}\right] \times \left[\frac{1}{2}, 1\right]$ for complements and $\left[\frac{1}{2}, 1\right] \times \left[0, \frac{1}{2}\right]$ for substitutes.

We have seen that norms can be strong enough to overcome material incentives and select many more equilibria than when norms are absent. For this reason, to discipline our analysis, we shall complement our model by modeling also the effect of actions on norms. It shall be the joint dynamics of norms and equilibrium actions to characterize the possible Nash equilibria depending on the strategic environment and the society composition.

1.2.3 Cultural Transmission

At the end of each time period t, given the action played during their adult ages, agents transmit new norms θ_{t+1} to their offsprings and decide how much let them socialize with the peers by transmitting an horizontal socialization f_{t+1}. In this section, we model both transmissions as a function of generation t norms, x_t, socializations, f_t, and average actions chosen when playing the one-stage games.

Formation of Transmitted Norms

First, we assume that agents try to reduce the *cognitive dissonance* that arises if there is no consistency between their original preference over actions x_t and their average behavior.

Regarding Nash equilibrium actions, agents are randomly matched to play against the whole population. Thus, an agent belonging to the majority, community $i = 1$, will be matched η time with his own type and $1 - \eta$ with others, viceversa for a player belonging to the minority. Moreover, as discussed in previous section, each agent in his adult life plays an infinite number of games with different payoffs $(\bar{b}, \bar{d})$ distributed according to a probability measure γ.

In order to get the average action played by the representative agent of one community we have to integrate actions with respect both population and payoffs measure, η and γ.[20] In particular, all the time agents play a game with a unique Nash equilibrium the transmitted action is uniquely defined to be the Nash equilibrium one. When, instead, the played game has not a unique Nash equilibrium, we assume indifference on the type of action that is transmitted. Thus each time an agent of community i plays against an agent of community j in period t,

[20]Notice that, due to the assumption of independence the order of integration does not affect the result and $E_{\eta,\gamma}[.] = E_{\gamma,\eta}[]$

the transmitted action is

$$A_{i,j,t} = \begin{cases} \hat{A}_{i,j,t}(\bar{b}, \bar{d}, x) & \text{if the Nash equilibrium is unique} \\ \frac{1}{2} & \text{otherwise} \end{cases} .$$

Whenever there are multiple equilibria, the feedback from actions to norms of a player of community i when playing with a player of community j in date t is $A_{i,j,t} = \frac{1}{2}$.[21]

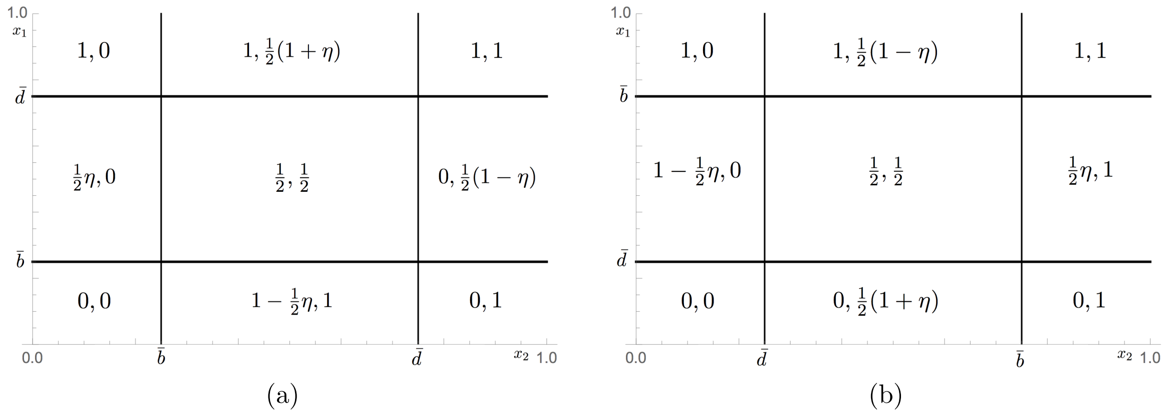


Figure 1.2: $E_\eta[A_{1,t}], E_\eta[A_{2,t}]$ in (a) Complements $\left(\bar{b} \leq \frac{1}{2} \leq \bar{d}\right)$, (b) Substitutes $\left(\bar{d} \leq \frac{1}{2} \leq \bar{b}\right)$ environments

Given a particular game, and $A_{i,j,t}$ for all $i, j = 1, 2$, the average action for both communities in period t are

$$\begin{cases} E_\eta[A_{1,t}] = \eta A_{1,1,t} + (1-\eta) A_{1,2,t} \\ E_\eta[A_{2,t}] = (1-\eta) A_{2,2,t} + \eta A_{2,1,t} \end{cases} \tag{1.5}$$

In Figure 1.2, we show average actions for agents of both communities for different (x_1, x_2), in both complements and substitutes game. It is easy to generalize (1.5) to take into account different levels of assortativity in the matching, in that case results do not change.

Considering all possibile games that agents face in their life and integrating with respect to the payoffs distribution γ, we finally get the vector of average actions:

$$E_{\eta,\gamma}[A_t] = (E_{\eta,\gamma}[A_{1,t}], E_{\eta,\gamma}[A_{2,t}]) = \varphi(x_t) \tag{1.6}$$

[21] A different possibility could be to consider the Mixed Nash equilibrium. In several empirical works the approximation of $\frac{1}{2}$ is widely used in presence of multiple equilibria (Bjorn and Vuong, 1984; Kooreman, 1994; Soetevent and Kooreman, 2007).

By *cognitive dissonance*, for a generic i we let the *ex ante* personal norms of generation $t+1$, θ_{t+1}, to move towards the transmitted actions $E_{\eta,\gamma}[A_t]$ as in

$$\theta_{i,t+1} = \zeta(x_{i,t}, E_{\eta,\gamma}[A_{i,t}]) = (1-\lambda)x_{i,t} + \lambda E_{\eta,\gamma}[A_{i,t}] \tag{1.7}$$

with $\lambda \in (0,1)$. If $\lambda = 0$ the strategic environment (games) has no effect on the evolution of preferences, this is the case considered by Kuran and Sandholm (2008). Otherwise, the inherited norm of the next generation depends directly on the norm of the parent x_t and indirectly also on the strategic environment through the expected equilibrium actions in $E_{\eta,\gamma}[A_t]$. Parents want to leave to their offsprings norms that are a combination of their norms and of what they have learnt to be the best action in the specific strategic environment they face.

Formation of Horizontal Socialization

The dynamics of the socialization level depends on the outcome of strategic environment and is modeled assuming *cultural substitution*.[22]

We assume that the more their action is close to the average of the society, the more agents let their offsprings horizontally socialize with the peers (the less they vertically socialize them). Given the average action in the whole society,

$$\bar{E}_{\eta,\gamma}[A_t] = \eta E_{\eta,\gamma}[A_{1,t}] + (1-\eta)E_{\eta,\gamma}[A_{2,t}],$$

the transmitted horizontal socialization for a generic agent of community i is

$$f_{i,t+1} = \psi(f_{i,t}, E_{\eta,\gamma}[A_t]) = \bar{f}(1 - |E_{\eta,\gamma}[A_{i,t}] - \bar{E}_{\eta,\gamma}[A_t]|). \tag{1.8}$$

The horizontal socialization of community i directly depends on the differences between the action of the agent and the action of the whole society. When the distance between actions is maximal, f goes to its lower bound. When the distance is minimal(agents play all the same action) f reaches the upper bound, $\bar{f} \leq 1$.
Notice that $\bar{f}$ is the maximum possible flexibility parameter of the society, namely the highest level of horizontal socialization of a community toward the whole society, and thus also to the other community.

[22]Under *cultural substitution* "parents have fewer incentives to socialize their children the more widely dominant are their values in the population" Bisin and Verdier (2001). Usually, in the literature of cultural transmission, cultural substitution is microfounded. Here we assume it. In Appendix B, we show that if parents optimally chose their socialization effort there is *cultural substitution*.

1.3 Norm and Socialization Level Dynamics

In this section, we consider the whole dynamics of the model and we present results for environments with *complements* and *substitutes*.

As anticipated in Section 1.2.1, in each period x depends directly on θ and f. Implicitly equation (1.1) for a generic i can be written as:

$$x_{i,t+1} = \upsilon(\theta_{i,t+1}, f_{i,t+1}). \tag{1.9}$$

Combining equations (1.6 - 1.9) we get the dynamics of our model

$$x_t \xrightarrow{\varphi(.)} E_{\eta,\gamma}[A_t] \xrightarrow{\zeta(.),\ \psi(.)} \theta_{t+1}, f_{t+1} \xrightarrow{\upsilon(.)} x_{t+1}.$$

The whole dynamics can be written as

$$\begin{cases} x_{t+1} = \Xi(x_t, f_t) \\ f_{t+1} = \Phi(x_t, f_t) \end{cases}, \tag{1.10}$$

where $\Xi(x_t, f_t) = \upsilon(\zeta(x_t, \varphi(x_t)), \psi(f_t, \varphi(x_t)))$ and $\Phi(x_t, f_t) = \psi(f_t, \varphi(x_t))$.

We name $\mathcal{E}$ the set of steady states of (1.10) with generic elements $(x_1^*, x_2^*, f_1^*, f_2^*)$. We further define $E_{\eta,\gamma}[A_i^*]$ as the average equilibrium action played by a representative agent of community $i \in \{1, 2\}$ at steady state, and $p_1^* = p_1(f_1^*, f_2^*)$ and $p_2^* = p_2(f_1^*, f_2^*)$, the equilibrium weights as defined in Corollary 1.1.

First, we provide a relation between norms and average actions at the steady state.

Proposition 3
Given the dynamics in (1.10), for all strategic environments, the norms (x_1^*, x_2^*) solve

$$\begin{cases} x_1^* = \phi_1^* E_{\eta,\gamma}[A_1^*] + (1 - \phi_1^*) E_{\eta,\gamma}[A_2^*] \\ x_2^* = \phi_2^* E_{\eta,\gamma}[A_1^*] + (1 - \phi_2^*) E_{\eta,\gamma}[A_2^*] \end{cases} \tag{1.11}$$

where $\phi_1^* = \frac{p_1^* - (p_1^* - p_2^*)(1-\lambda)}{1-(p_1^* - p_2^*)(1-\lambda)}$.

Proof. In the Appendix. □

In a steady state, norms are a convex combination of the average actions played by the agents of the two communities. Weights depend both on the steady state horizontal socialization levels f^*, through weight p_1^* and p_2^*, and on two exogenous variables, the size of the majority η and the influence of games on norms λ.

We now present steady state norms, horizontal socialization, and actions' vector played by agents of the two communities for both *complements* and *substitutes* environment. We consider both uniform and point payoffs' distributions.

1.3.1 Uniform Distribution

In this section γ is a uniform distribution on sets of $(\bar{b}, \bar{d})$. We divide results between *complements* and *substitutes*.

If in adult age agents face a *complements* environment they play several games in which $\bar{b} \leq \frac{1}{2} \leq \bar{d}$.

Below, we characterize the set of possible steady states together with their global stability.

Proposition 4 (Complements)
For all $\eta \in (0,1)$, $\bar{f} \in (0,1)$, and $\lambda \in (0,1)$, if $(\bar{b}, \bar{d})$ is uniformly distributed in the set $\left[0, \frac{1}{2}\right] \times \left[\frac{1}{2}, 1\right]$, then

- *the set of steady states is*

$$\mathcal{E} = \{(1,1,\bar{f},\bar{f}), (0,0,\bar{f},\bar{f}), \left(\frac{1}{2}, \frac{1}{2}, \bar{f}, \bar{f}\right)\};$$

- *the average actions at the steady state, $E_{\eta,\gamma}[A^*]$, are, respectively,*

$$(1,1), (0,0), \left(\frac{1}{2}, \frac{1}{2}\right);$$

- *for all initial conditions the dynamics converges to an element of the set $\mathcal{E}$. Moreover, the basin of attraction of $(1,1,\bar{f},\bar{f})$ is at least $\left[\frac{1}{2}, 1\right]^2 \times \left[0, \bar{f}\right]^2$, the basin of attraction of $(0,0,\bar{f},\bar{f})$ is at least $\left[0, \frac{1}{2}\right]^2 \times \left[0, \bar{f}\right]^2$, and $\left(\frac{1}{2}, \frac{1}{2}, \bar{f}, \bar{f}\right)$ is a saddle.*

Proof. in the Appendix □

In a *complements* environment, if payoffs are uniformly distributed over all the support $\left[0, \frac{1}{2}\right] \times$

$\left[\frac{1}{2}, 1\right]$, the only possible long-run outcome are assimilation toward one extreme norm. In fact, both communities always converge to the same norms and the horizontal socialization is at its maximum. In this kind of equilibria norms and action played at the steady state are the same.

We now discuss results for a *substitutes* environment, where agents play several games in which $\bar{d} \leq \frac{1}{2} \leq \bar{b}$.

Proposition 5 (Substitutes)
For all $\eta \in (0,1)$, $\bar{f} \in (0,1)$, and $\lambda \in (0,1)$, if $(\bar{b}, \bar{d})$ is uniformly distributed in the set $\left[0, \frac{1}{2}\right] \times \left[\frac{1}{2}, 1\right]$, then

- *the set of steady states is*

$$\mathcal{E} = \{(1,1,\bar{f},\bar{f}), (0,0,\bar{f},\bar{f}), \left(\frac{1}{2}, \frac{1}{2}, \bar{f}, \bar{f}\right), (\phi_1, \phi_2, \bar{f}\eta, \bar{f}(1-\eta)), (1-\phi_1, 1-\phi_2, \bar{f}\eta, \bar{f}(1-\eta))\};$$

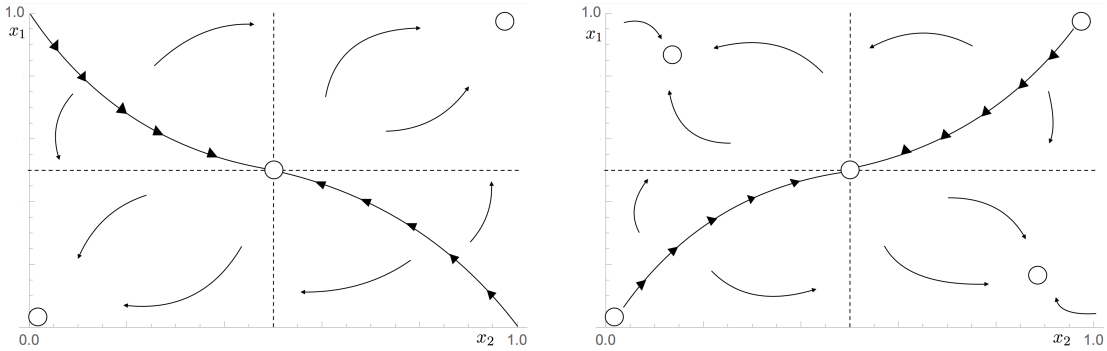
- *the average actions at the steady state, $E_{\eta,\gamma}[A^*]$, are, respectively*

$$(1,1), (0,0), \left(\frac{1}{2}, \frac{1}{2}\right), (1,0), (0,1);$$

- *For all initial conditions the dynamics converges to an element of the set. Moreover, $(1,1,\bar{f},\bar{f})$ and $(0,0,\bar{f},\bar{f})$ are unstable, the basin of attraction of $(\phi_1, \phi_2, \bar{f}\eta, \bar{f}(1-\eta))$ and $(1-\phi_1, 1-\phi_2, \bar{f}\eta, \bar{f}(1-\eta))$ is at least $\left[\frac{1}{2}, 1\right] \times \left[0, \frac{1}{2}\right] \times \left[0, \bar{f}\right]^2$ and $\left[0, \frac{1}{2}\right] \times \left[\frac{1}{2}, 1\right] \times \left[0, \bar{f}\right]^2$, respectively, and $\left(\frac{1}{2}, \frac{1}{2}, \bar{f}, \bar{f}\right)$ is a saddle.*

Proof. In the Appendix □

In an environment with *substitutes*, if payoffs are uniformly distributed over all the support $\left[\frac{1}{2}, 1\right] \times \left[0, \frac{1}{2}\right]$, there still exist steady states with *assimilation* as in the previous (*complements*) case, however they are unstable. Moreover, there exist steady states with separation. In the latter norms are polarized ($\phi_1 > \frac{1}{2}$ and $\phi_2 < \frac{1}{2}$), horizontal socialization is at its minimum for both communities ($\eta\bar{f}, (1-\eta)\bar{f}$) and agents of the two communities always play opposite actions $E_{\eta,\gamma}[A^*] = (0,1)$ or $E_{\eta,\gamma}[A^*] = (1,0)$. In these cases the minority tends to be culturally separated, in fact the larger is the size of the majority η the smaller the horizontal socialization level of the minority becomes. Here the effect of coordination game played by children is evident. If they care only about inherited norms θ, because parents are able to fully vertically socialize the offspring ($f_1 = f_2 = 0$), then norms became fully polarized ($\phi_1 = 1$ and $\phi_2 = 0$).



(a) (b)

Figure 1.3: Qualitative dynamics for (a) Complements and (b) Substitutes environments

The main implication of Propositions 4 and 5 is that different environments lead to different social outcomes. In particular, we have shown how in *substitutes* environments emerge steady states with polarization of norms that do not exist in *complements* ones. Moreover, although steady states with assimilation may exist in both cases, when games played in adult age are *substitutes*, this kind of steady states are not stable. It is important to point out that, according to Proposition 4, an environment with *complements* always lead to norms homogeneity.

1.3.2 Point Distribution

Despite our results, an empirical observation is that an environment that favors coordination is not always enough to ensure the occurrence of complete assimilation or to avoid the occurrence of polarization in society. There are both cases in which the process of cultural integration occurs without achieving complete assimilation, having the resilience of cultural traits, as well as cases of cultural separation and the rise of oppositional cultures.[23] Bisin and Verdier (2012) offer a review of empirical examples of cultural heterogeneity and resilience of cultural traits. For example, the slow rate of immigrants' integration in Europe and US, the persistence of 'ethnic capital' in second- and third-generation immigrants, or even cases of minorities' strongly attachment to languages and cultural traits.

In this section, we show how to reconcile our model of norm formation with these empirical observation. We do so by considering limit cases of point distribution namely when γ is

[23]We refer to Berry (1997) and Ryder et al. (2000) (among others) for the terminology about cultural *assimilation*, *integration*, *marginalization* and *separation*. They proposed a concept of minority's self-identification, based on a two-dimensional framework, which takes into account for differences in both adaptation and interaction processes between the minority and the dominant culture.

singular on the set $(\bar{b}, \bar{d})$ showing how polarization can occur even in environments with *complements* and milder cultural heterogeneity can occur in environments with both *complements* and *substitutes*.[24]

Before we discuss the possible steady states in both *complements* and *substitutes* environment, we first provide a general result.

Proposition 6 *For all* $\eta \in (0,1)$*,* $\bar{f} \in (0,1)$*, and* $\lambda \in (0,1)$*, if* $(\bar{b}, \bar{d})$ *is a point in the set* $[0,1] \times [0,1]$*, then all steady states are asymptotically stable and their basins of attraction form a partition of the state space of norms and socialization levels.*

Proof. In the Appendix. □

Proposition 6 ensures the convergence and the stability of dynamics described in (1.10). That is, for every initial conditions and for all parameters, the dynamics converges to a steady state norm and flexibility parameter. The main idea is that we can partition the whole state space of norms and socialization levels, $[0,1]^2 \times [0,1]^2$, in basins of attraction of steady states.

Complements

Figure 1.4 shows all the possible steady states in environments with *complements*.

The first result is that two stable steady states with cultural separation (polarized norms and minimal horizontal socialization) can exist depending on the value of $\bar{b}$ and $\bar{d}$ and on the norms' distance between the two communities. For simplicity, we now formally present only one of the two steady states with cultural separation (blu in Figure 1.4), the other can be easily derived by symmetry.

If $\bar{b} > \phi_1^*$ and $\bar{d} < \phi_2^*$ and initial norms belong to the region $R_{1_*,0_*}$, then the dynamics described in (1.10) converges to the steady state $(\phi_1^*, \phi_2^*, \bar{f}\eta, \bar{f}(1-\eta))$, where norms are polarized, horizontal socialization is at the minimum, and agents belonging to different communities always play different actions, $E_{\eta,\gamma}[A^*] = (1,0)$. This occurs despite in all rounds agents are playing a coordination game, according to material payoffs.

The result is driven from the fact that if games played in adult life have always the same non degenerate payoffs, then there exist initial norms, x_1 high enough and x_2 low enough, that

[24] We present and discuss the results informally, we remand to the appendix to have formal derivation and discussion.

sustain the equilibrium with anti-coordination. Playing this equilibrium leads, by *cognitive dissonance*, to norms polarization and thus cultural *separation*.

If agents interact in *complements* another class of steady states can exist (white in Figure 1.4), where there is cultural *integration* but not complete *assimilation*. In such steady states, only the agents of one community have a well-defined group-specific norm, which induce them to always play a specific action as dominant strategy. While agents belonging to the other community have a mild norm. When the latter are matched among themselves, they face the original coordination problem, while they conform to the behavior of agents belonging to the other community whenever they encounter them. This is a clear example of *integration*, namely there is convergence toward an homogeneous norms, but, at the same time, the identity is not totally lost as in *assimilation*.

Real life examples of this result are linguistic choices between immigrants and natives. Natives always use their own language. Agents belonging to a linguistic minority, after a long interaction with natives, became proficient in both languages (i.e. second or third generation immigrants). Therefore, they end up using the two languages indifferently, but conforming with the natives whenever they interact with them.

In steady states with *integration*, we have two sources of symmetry. One is with respect to the community with a well defined norm, the other with respect to the action played. Therefore there can exist up to four steady states of this type. Again we formally characterize only one of these equilibria.

E.g. let us focus on the region $R_{1_*,1}$, were there can exist a steady state where $\bar{b} < x_2 < \bar{d}$ and $E_{\eta,\gamma}[A] = \left(1, \frac{1}{2}(1+\eta)\right)$.

Notice that with point distribution this kind of steady state are stable in their basin of attraction, but may do not exists for certain values of $(\bar{b}, \bar{d})$ as shown in Figure 1.5. While steady states with assimilation always exist, those with integration may not exist (see for example Figure 1.5).

The difference between *assimilation* and *integration* equilibria is important with respect to different policy goals. For example, sometimes a policy is considered successful only when minorities (immigrants) completely lose their previous norms or culture and are assimilated; instead in other circumstances the resilience of cultural traits can be considerate socially desirable, in these cases the policymaker reaches his goal if the minority integrate with the majority

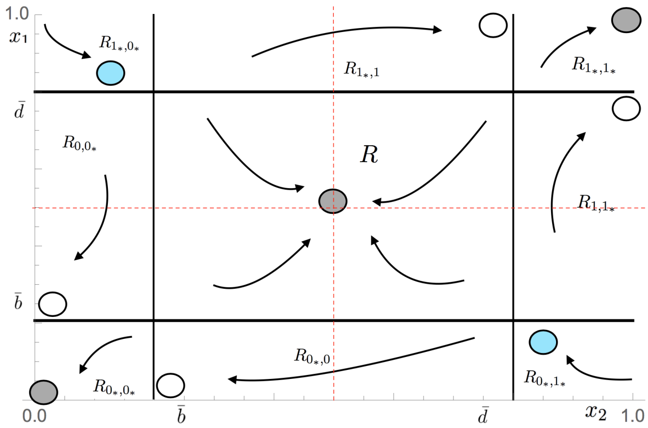


Figure 1.4: Steady States of x for *Complements* with $\bar{b} = 0.2$, $\bar{d} = 0.8$, $\bar{f} = 0.3$, $\eta = 0.5$, $\lambda = 0.6$.

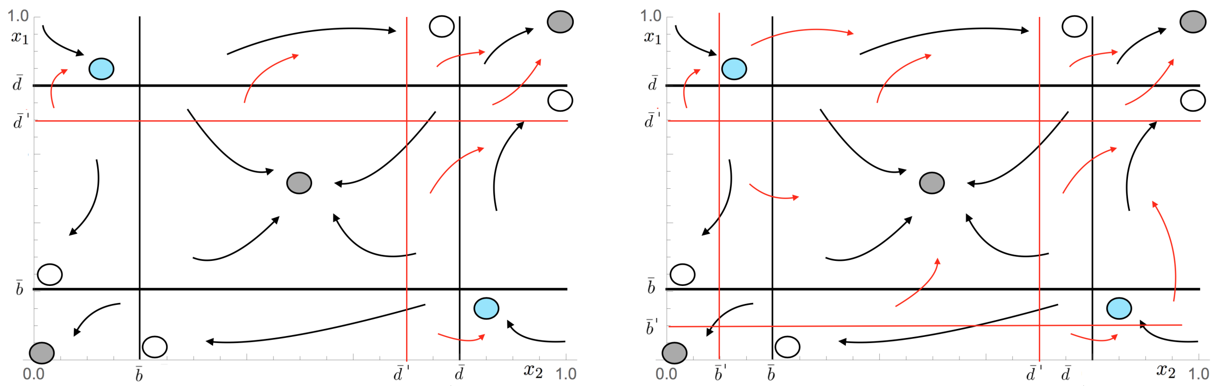


Figure 1.5: Steady States of x for *Complements* with $\bar{b} = 0.2$, $\bar{d} = 0.8$, $\bar{d}' = 0.7$, $\bar{b}' = 0.1$, $\bar{f} = 0.3$, $\eta = 0.5$, $\lambda = 0.6$.

but keeping some of their cultural traits. In this second case, there is a partial convergence and there is still room for a multicultural society.

Moreover, from a policy maker point of view, it is interesting to appraise the long-run effects due to a change in game incentives (Figure 1.5). We observe that moving the parameters $\bar{b}, \bar{d}$ can significantly affect the social outcome. In Figure 1.5 (left panel) we can see how diminishing $\bar{d}$ to $\bar{d}'$ the basin of attraction of the steady state with assimilation (grey) becomes much wider and integration (white) disappears. Figure 1.5 (right panel) shows that in such a case, moving the two material incentives together, $\bar{d}$ to $\bar{d}'$ and $\bar{b}$ to $\bar{b}'$, it is possible to reach assimilation even if the communities start off having initial norms that are polarized. This sheds further light on the relationship between uniform and point distribution. In fact, if $(\bar{b}, \bar{d})$ is not a single point, but moves in the whole space $\left[0, \frac{1}{2}\right] \times \left[\frac{1}{2}, 1\right]$, then the steady states presented in this section are not robust to game change.

Substitutes

Figure 1.6 shows all the possible steady states in *substitutes* environments when γ is singular. The main difference with the case of uniform distribution is that steady states with assimilation can be stable. Moreover, in environments with *substitutes*, steady states may exist (white in Figure 1.6) where one community has a well defined norm while the members of the other community have a norm that, when they are matched among themselves, does not induce preferences over actions. Namely, interacting in their own community agents are indifferent on the action to play, but when matched with the other community they act in the opposite way. In a sense, we can talk about *marginalization*, in fact, there is a weakening of identification with original cultural identity but even a rejection of dominant culture.

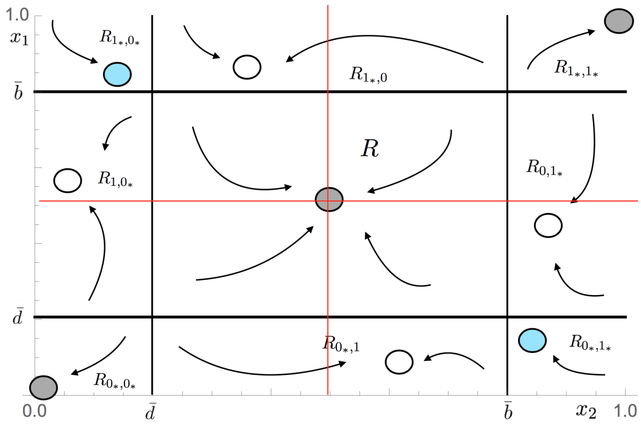


Figure 1.6: Steady States of x for *Substitutes* with $\bar{b} = 0.8$, $\bar{d} = 0.2$, $\bar{f} = \eta = 0.5$, $\lambda = 0.75$.

1.4 Discussion

Payoffs Distribution This paper is a first attempt to analyze the effect of different environment on norms that arise in the society. We study only extreme cases for the payoffs' distribution (uniform and point). We can easily generalize our results for all possible uniform distributions. The relevant parameters to understand the existence of different steady states and the dynamics are the boundary points (the highest and the lower) of the support of the distribution and their relationship with the material incentives $\bar{b}, \bar{d}$. Namely, in the interior points of the support the dynamics is the same as studied in 3.1, while outside it results discussed in 3.2 hold. Considering probability distribution functions different from the uniform one, the analysis is less straightforward, it mainly depends on the density in the tails. Our conjecture is that the relevant parameters are threshold values over which the density is vanishing, and thus the probability that an agent face that particular game is extremely low. Thresholds should depends, non trivially, on the second, third and fourth moments. Although all these possible intermediate cases could can be interesting to study from a technical point of view, we believe that they should not provide any novel qualitative insight.

Assortativity We have studied the case of perfect random matching without taking into account the possibility of having assortative matching. In order to consider different levels of assoratativity it is enough to consider a paramenter ε that can assume values less then $1-\eta$ for the majority and η for the minority and add to the probability of being matched with agents belonging to the same community. This generalization does not affect results, and the proof of all the proposition remain the same.[25] The only effect of an higher level of assortativity is to slow down the convergence to the steady state, and thus the assimilation of norms, in environment with *complements* and increasing the speed of convergence to the steady state, and thus the polarization, in environment with *substitutes*.

Mixed Environments While in this paper we keep the type of environment (*complements* or *substitutes*) fixed, an extension of the model allowing changes in the type of environment should be worth to be investigated. Our conjecture is that norms may not converge and generate cycles.The same results should be found also when the strategic environment resembles Prisoner Dilemma. In fact, according to our preliminary analysis, depending on the material payoffs parametrization, the Prisoner Dilemma can act both as *complements* or *substitutes* environment. This extension is particularly relevant to study more complex environments where some interactions are cooperative and other competitive. Moreover, our analysis suggests that if in a society there are many tasks that require a joint effort, then agents develop more cooperative norms. This is supported by the empirical evidence that suggests how tough environments with less developed institutions, or fewer resources, favor norms of cooperation Lowes et al. (2017). Since often institutions (*complements substitutes*) are endogenous and co-evolve within the society, both because agents have, in some circumstances, the chance to vote to choose their institution, or because policy maker can design different institutions depending on the society, can create non trivial culture-institution dynamics.

1.5 Conclusion

In this paper, we study a cultural transmission model where the relationship between norms and strategic environments is made explicit. Agents divided into two communities form their community norm by taking into account the norm received by their previous generation and by conforming to the average norm of the society. The relative strength of the two forces is

[25]The only difference is to have $\eta' = \eta + \epsilon$ instead of η, but this does not play a role, the only condition relevant for the proofs is that $\eta' < 1$ and, thus, $1 - \eta' > 0$.

regulated by a horizontal socialization parameter. The norm received by the previous generation depends on the average equilibrium actions played in the game under the hypothesis of minimization of *cognitive dissonance*. We derive conditions under which cultural assimilation is reachable or not. Provided games material payoffs are randomly distributed but preserve their *complements/substitutes* (coordination/anti-coordination) feature, the long-run dynamics converges to cultural assimilation in environments with *complements* and to cultural segregation in environment with *substitution*. Moreover, when specific games are chosen, provided initial conditions show enough heterogeneity, we are able to obtain the rise of oppositional cultures and situations of cultural heterogeneity. For example, we show that even if the environment provides incentives to coordinate it is still possible to observe multicultural society or even cultural separation. At the same time in anti-coordinating environments it is still possible to reach assimilation.

Appendix A

Proof of Proposition 1

In order to simplify the notation, we consider payoffs for a representative agent of each community. The payoff of an generic agent belonging to community $i = 1, 2$

$$u_i(x, \theta_i, f_i) = -f_i \left(x_i - \underbrace{(\eta x_1 + (1-\eta)x_2)}_{E_\eta[x]} \right)^2 - (1 - f_i)(x_i - \theta_i)^2$$

becomes Since the single agent is negligible in the population, when i takes decisions he does not affect the whole average. Thus, the first order condition is

$$\frac{\partial u_i(x, \theta_i)}{\partial x_i} = 0$$

$$-2f_i\left(x_i - (\eta x_1 + (1-\eta)x_2)\right) - 2(1 - f_i)(x_i - \theta_i) = 0$$

$$f_i\left(x_i - (\eta x_1 + (1-\eta)x_2)\right) + (1 - f_i)(x_i - \theta_i) = 0.$$

As a result

$$x_i = f_i E_\eta[x] + (1 - f_i)\theta_i \tag{1.12}$$

Taking expectations of (1.12) on both sides we get

$$E_\eta\left[x\right] = E_\eta\left[f E_\eta[x] + (1-f)\theta\right]$$

$$E_\eta\left[x\right] = E_\eta\left[f E_\eta[x]\right] + E_\eta\left[\theta\right] - E_\eta[f\theta]$$

$$E_\eta\left[x\right] - E_\eta\left[f\right] E_\eta[x] = E_\eta\left[\theta\right] - E_\eta[f\theta]$$

$$E_\eta\left[x\right] - E_\eta\left[f\right] E_\eta[x] = E_\eta\left[\theta\right] - E_\eta[f] E_\eta[\theta] - cov_\eta[f, \theta]$$

$$(1 - E_\eta[f]) E_\eta\left[x\right] = (1 - E_\eta[f]) E_\eta\left[\theta\right] - cov[f, \theta]$$

$$E_\eta\left[x\right] = E_\eta\left[\theta\right] - \frac{cov_\eta[f, \theta]}{(1 - E_\eta[f])}$$

.

Substituting $E_\eta[x]$ in (12) we find the optimal action of each player belonging to community i

as a function of the distributions of θ and f

$$x_i = f_i \left(E_\eta [\theta] - \frac{cov_\eta[f,\theta]}{(1 - E_\eta[f])} \right) + (1 - f_i)\theta_i \tag{1.13}$$

□

Proof of Corollary 1.1

Applying the ex-post norm formation rule (1.13) to our particular case where η is the share of agents with ex-ante norm θ_1 and $1 - \eta$ is the share of agents with θ_2, we get:

$$x_1 = f_1 \left(\eta\theta_1 + (1 - \eta)\theta_2 - \frac{cov_\eta[f,\theta]}{1 - f_1\eta - f_2(1 - \eta)} \right) + (1 - f_1)\theta_1$$

Computing

$$cov_\eta[f,\theta] = E_\eta [(f - E_\eta[f])(\theta - E_\eta[\theta])] =$$

$$= \eta(f_1 - f_1\eta - f_2(1-\eta))(\theta_1 - \theta_1 n - \theta_2(1-\eta)) + (1-n)(f_2 - f_1\eta - f_2(1-\eta))(\theta_2 - \theta_1\eta - \theta_2(1-\eta))$$

$$= \eta(1 - \eta)^2(f_1 - f_2)(\theta_1 - \theta_2) + \eta^2(1 - \eta)(f_2 - f_1)(\theta_2 - \theta_1)$$

$$= \eta(1 - \eta)(f_1 - f_2)(\theta_1 - \theta_2)((1 - \eta) + \eta)$$

$$\Rightarrow cov_\eta[f,\theta] = \eta(1 - \eta)(f_1 - f_2)(\theta_1 - \theta_2)$$

we obtain

$$x_1 = f_1 \left(\eta\theta_1 + (1 - \eta)\theta_2 - \frac{\eta(1 - \eta)(f_1 - f_2)(\theta_1 - \theta_2)}{1 - f_1\eta - f_2(1 - \eta)} \right) + (1 - f_1)\theta_1$$

$$x_1 = \underbrace{\frac{f_1(1 - \eta)(1 - f_2)}{1 - f_1\eta - f_2(1 - \eta)}}_{1-p_1} \theta_2 + \underbrace{\frac{(1 - f_1)(1 - f_2(1 - \eta))}{1 - f_1\eta - f_2(1 - \eta)}}_{p_1} \theta_1$$

The same can be done for x_2.

□

Corollary 2.1

Before to proceed with other proofs we formally partition the space $[0, 1]^2$ in regions where the different Nash equilibria emerge as described in Figure 3.2.

Corollary 2.1 The regions of norms x_{i_r} and x_{j_c} in which different Nash Equilibria emerge are the following:

$$\begin{cases} R_{1_*,1_*} = \{(x_{i_r}, x_{j_c}) : x_{i_r} > max\{\bar{d}, \bar{b}\} \wedge x_{j_c} > max\{\bar{d}, \bar{b}\}\} \\ R_{0_*,0_*} = \{(x_{i_r}, x_{j_c}) : x_{i_r} < min\{\bar{d}, \bar{b}\} \wedge x_{j_c} < min\{\bar{d}, \bar{b}\}\} \\ R = \{(x_n, x_m) : min\{\bar{d}, \bar{b}\} < x_{i_n} < max\{\bar{d}, \bar{b}\} \wedge min\{\bar{d}, \bar{b}\} < x_{i_m} < max\{\bar{d}, \bar{b}\}\} \\ R_{1_*,0_*} = \{(x_{i_r}, x_{j_c}) : x_{i_r} > max\{\bar{d}, \bar{b}\} \wedge x_{j_c} < min\{\bar{d}, \bar{b}\}\} \\ R_{0_*,1_*} = \{(x_{i_r}, x_{j_c}) : x_{i_r} < min\{\bar{d}, \bar{b}\} \wedge x_{j_c} > max\{\bar{d}, \bar{b}\}\} \\ R_{1_*,1} = \{(x_{i_r}, x_{j_c}) : x_{i_r} > max\{\bar{d}, \bar{b}\} \wedge min\{\bar{d}, \bar{b}\} < x_{j_c} < max\{\bar{d}, \bar{b}\} \wedge \bar{d} > \bar{b}\} \\ R_{1_*,0} = \{(x_{i_r}, x_{j_c}) : x_{i_r} > max\{\bar{d}, \bar{b}\} \wedge min\{\bar{d}, \bar{b}\} < x_{j_c} < max\{\bar{d}, \bar{b}\} \wedge \bar{b} > \bar{d}\} \\ R_{1,1_*} = \{(x_{i_r}, x_{j_c}) : x_{j_c} > max\{\bar{d}, \bar{b}\} \wedge min\{\bar{d}, \bar{b}\} < x_{i_r} < max\{\bar{d}, \bar{b}\} \wedge \bar{d} > \bar{b}\} \\ R_{0,1_*} = \{(x_{i_r}, x_{j_c}) : x_{j_c} > max\{\bar{d}, \bar{b}\} \wedge min\{\bar{d}, \bar{b}\} < x_{i_r} < max\{\bar{d}, \bar{b}\} \wedge \bar{b} > \bar{d}\} \\ R_{0_*,0} = \{(x_{i_r}, x_{j_c}) : x_{i_r} < min\{\bar{d}, \bar{b}\} \wedge min\{\bar{d}, \bar{b}\} < x_{j_c} < max\{\bar{d}, \bar{b}\} \wedge \bar{d} > \bar{b}\} \\ R_{0_*,1} = \{(x_{i_r}, x_{j_c}) : x_{i_r} < min\{\bar{d}, \bar{b}\} \wedge min\{\bar{d}, \bar{b}\} < x_{j_c} < max\{\bar{d}, \bar{b}\} \wedge \bar{b} > \bar{d}\} \\ R_{0,0_*} = \{(x_{i_r}, x_{j_c}) : x_{i_r} < min\{\bar{d}, \bar{b}\} \wedge min\{\bar{d}, \bar{b}\} < x_{j_c} < max\{\bar{d}, \bar{b}\} \wedge \bar{d} > \bar{b}\} \\ R_{1,0_*} = \{(x_{i_r}, x_{j_c}) : x_{j_c} < min\{\bar{d}, \bar{b}\} \wedge min\{\bar{d}, \bar{b}\} < x_{i_r} < max\{\bar{d}, \bar{b}\} \wedge \bar{b} > \bar{d}\} \end{cases}$$

Proof of Proposition 3

Substituting the dynamics of θs in equation (1.3) we get

$$\begin{cases} x_{1,t+1} = p_{1,t+1}[(1-\lambda)x_{1,t} + \lambda E_{\eta,\gamma}[A_{1,t}]] + (1-p_{1,t+1})[(1-\lambda)x_{2,t} + \lambda E_{\eta,\gamma}[A_{2,t}]] \\ x_{2,t+1} = p_{2,t+1}[(1-\lambda)x_{1,t} + \lambda E_{\eta,\gamma}[A_{1,t}]] + (1-p_{2,t+1})[(1-\lambda)x_{2,t} + \lambda E_{\eta,\gamma}[A_{2,t}]] \end{cases} \tag{1.14}$$

$$\Rightarrow \begin{cases} x_{1,t+1} = \frac{p_{1,t+1}\lambda}{1-p_{1,t+1}(1-\lambda)} E_{\eta,\gamma}[A_{1,t}] + \frac{(1-p_{1,t+1})\lambda}{1-p_{1,t+1}(1-\lambda)} E_{\eta,\gamma}[A_{2,t}] + \frac{(1-p_{1,t+1})\lambda}{1-p_{1,t+1}(1-\lambda)} x_{2,t} \\ x_{2,t+1} = \frac{(1-p_{2,t+1})\lambda}{1-(1-p_{2,t+1})(1-\lambda)} E_{\eta,\gamma}[A_{2,t}] + \frac{p_{2,t+1}\lambda}{1-(1-p_{2,t+1})(1-\lambda)} E_{\eta,\gamma}[A_{1,t}] + \frac{p_{2,t+1}\lambda}{1-(1-p_{2,t+1})(1-\lambda)} x_{1,t} \end{cases}$$

Where $p_{1,t} = \frac{(1-f_{1,t})(1-f_{2,t}(1-\eta))}{1-f_{1,t}\eta - f_{2,t}(1-\eta)}$, $p_{2,t} = \frac{f_{2,t}\eta(1-f_{1,t})}{1-f_{1,t}\eta - f_{2,t}(1-\eta)}$.

Substituting and computing the steady state we get

$$\Rightarrow \begin{cases} x_1^* = \phi_1^* E_{\eta,\gamma}[A_1^*] + (1-\phi_1) E_{\eta,\gamma}[A_2^*] \\ x_2^* = \phi_2^* E_{\eta,\gamma}[A_1^*] + (1-\phi_2) E_{\eta,\gamma}[A_2^*] \end{cases} \tag{1.15}$$

Where $\phi_1^* = \frac{p_1^*-(p_1^*-p_2^*)(1-\lambda)}{1-(p_1^*-p_2^*)(1-\lambda)}$ and $\phi_2^* = \frac{p_2^*}{1-(p_1^*-p_2^*)(1-\lambda)}$.
Substituting p_1^* and p_2^* we get ϕ_1^*, ϕ_2^* depending only on $f_1^*, f_2^*, \eta, \lambda$

$$\begin{cases} \phi_1^* = \frac{(1-f_1^*)f_2\eta+\lambda(1-f_1^*-f_2^*+f_1^*f_2^*)}{f_1(1-\eta)+f_2^*\eta-f_1^*f_2^*+\lambda(1-f_1^*-f_2^*+f_1^*f_2^*)} \\ \phi_2^* = \frac{f_2^*\eta(1-f_1)}{f_1(1-\eta)+f_2^*\eta-f_1^*f_2^*+\lambda(1-f_1^*-f_2^*+f_1^*f_2^*)} \end{cases}$$

□

Proof of Proposition 4

Let start to define the dynamics for horizontal socialization

$$\begin{cases} f_1 = \bar{f}(1-(1-\eta)|(E_{\eta,\gamma}[A_{1,t}] - E_{\eta,\gamma}[A_{2,t}])|) \\ f_2 = \bar{f}(1-\eta|(E_{\eta,\gamma}[A_{2,t}] - E_{\eta,\gamma}[A_{1,t}]])|) \end{cases} \tag{1.16}$$

To verify that $(1,1,\bar{f},\bar{f}),(0,0,\bar{f},\bar{f})$ and $\left(\frac{1}{2},\frac{1}{2},,\bar{f},\bar{f}\right)$ belong to the set of steady states is enough to substitute them in (1.14) and (1.16)

To understand the dynamics we need to study $E_{\eta,\gamma}[A_1^*]$ and $E_{\eta,\gamma}[A_2^*]$ is the space $[0,1]^2$. In fact $E_{\eta,\gamma}[A_1^*]$ and $E_{\eta,\gamma}[A_2^*]$ affect directly the horizontal socialization dynamics (1.8) and the norms dynamics through (1.7) and (1.8). By symmetry we can focus only on regions $\{(x_1,x_2) : x_1 > \frac{1}{2}, x_2 > \frac{1}{2}\}$ and $\{(x_1,x_2) : x_1 > \frac{1}{2}, x_2 < \frac{1}{2}\}$.

- If $x_1 > \frac{1}{2}$ and $x_2 > \frac{1}{2}$

  $$E_{\eta,\gamma}[\bar{A}_1] = \sum A_1^i p(A_1 = A_1^i) = 1 * \gamma(\bar{d} < x_1) + \left(1 - \frac{1}{2}\eta\right) * \gamma(\bar{d} > x_1 \wedge \bar{d} < x_2) + \frac{1}{2} * \gamma(\bar{d} > x_1 \wedge \bar{d} > x_2)$$

  Where

  –
  $$\gamma(\bar{d} < x_1) = \int_{1/2}^{x_1} \gamma(\bar{d}) d\bar{d} = \int_{1/2}^{x_1} 2 d\bar{d} = 2\left(x_1 - \frac{1}{2}\right)$$

  –
  $$\gamma(\bar{d} > x_1 \wedge \bar{d} < x_2) = 4(1-x_1)\left(x_2 - \frac{1}{2}\right)$$

  –
  $$\gamma(\bar{d} > x_1 \wedge \bar{d} > x_2) = 4(1-x_1)(1-x_2)$$

Therefore

$$E_{\eta,\gamma}[A_1] = 2\left(x_1 - \frac{1}{2}\right) + \left(1 - \frac{1}{2}\eta\right) * 4\left(1 - x_1\right)\left(x_2 - \frac{1}{2}\right) + \frac{1}{2} * 4\left(1 - x_1\right)\left(1 - x_2\right)$$

$$= 2x_1 - 1 + \left(1 - \frac{1}{2}\eta\right) * 2\left(1 - x_1\right)\left(2x_2 - 1\right) + 2\left(1 - x_1\right)\left(1 - x_2\right)$$

$$= 2x_1 - 1 - 2 + 2x_1 + 4x_2 - 4x_1x_2 + (1 - x_1 - 2x_2 + 2x_1x_2)\eta + 2 - 2x_1 - 2x_2 + 2x_1x_2$$

$$= -1 + 2x_1 + 2x_2 - 2x_1x_2 + (1 - x_1 - 2x_2 + 2x_1x_2)\eta$$

When x_1 and x_2 are both larger than $\frac{1}{2}$, if $E_{\eta,\gamma}[A_1] > x_1$, then by symmetry $E_{\eta,\gamma}[A_2] > x_2$. Therefore, since $x_{1,t}$ and $x_{2,t}$ are a convex combination of $\theta_{1,t}$ and $\theta_{2,t}$ (3) and $\theta_{1,t}$ and $\theta_{2,t}$ grow if $E_{\eta,\gamma}[A_1] > x_1$ and $E_{\eta,\gamma}[A_2] > x_2$ respectively, we can conclude that $x_{1,t}$ and $x_{2,t}$ grow whenever $E_{\eta,\gamma}[A_1] > x_1$ and $E_{\eta,\gamma}[A_2] > x_2$. Therefore, to prove that $(1, 1, \bar{f}, \bar{f})$ is the only steady state of (10), and that the region $\{(x_1, x_2) : x_1 > \frac{1}{2}, x_2 > \frac{1}{2}\}$ has no cycle and that it is always the basin of attraction of $(1, 1, \bar{f}, \bar{f})$ it is enough to prove that $E_{\eta,\gamma}[A_1] > x_1$ and $E_{\eta,\gamma}[A_2] > x_2$. By symmetry, the same reasoning applies to $(0, 0, \bar{f}, \bar{f})$.

Let us verify that$E_{\eta,\gamma}[A_1] \geq x_1$

$$-1 + 2x_1 + 2x_2 - 2x_1x_2 + (1 - x_1 - 2x_2 + 2x_1x_2)\eta \geq x_1$$

$$-1 + x_1 + 2x_2 - 2x_1x_2 + (1 - x_1 - 2x_2 + 2x_1x_2)\eta \geq 0$$

$$\underbrace{(x_1 - 1)}_{\leq 0}\underbrace{(1 - 2x_2)}_{\leq 0}\underbrace{(1 - \eta)}_{>0} \geq 0 \qquad \textit{always satisfied}$$

$$\Rightarrow E_{\eta,\gamma}[A_1] \geq x_1 \Rightarrow x_{t,1} - x_{t-1,1} > 0$$

By symmetry

$$\Rightarrow E_{\eta,\gamma}[A_2] \geq x_2 \Rightarrow x_{t,2} - x_{t-1,2} > 0$$

- If $x_1 > \frac{1}{2}$ and $x_2 < \frac{1}{2}$

$$E_{\eta,\gamma}[A_1] = 1 * \gamma(\bar{d} < x_1) + \frac{1}{2}\eta * \gamma(\bar{d} > x_1 \ \wedge \ \bar{b} > x_2) + \frac{1}{2} * \gamma(\bar{d} > x_1 \wedge \bar{b} < x_2)$$

Where

–

$$\gamma(\bar{d} < x_1) = \int_{1/2}^{x_1} \gamma(\bar{d})d\bar{d} = \int_{1/2}^{x_1} 2d\bar{d} = 2\left(x_1 - \frac{1}{2}\right)$$

–

$$\gamma(\bar{d} > x_1 \ \wedge \ \bar{b} > x_2) = \left(1 - \int_{1/2}^{x_1} 2d\bar{d}\right) * \left(1 - \int_{0}^{x_2} 2d\bar{b}\right) = 2(1-x_1)(1-2x_2)$$

–

$$\gamma(\bar{d} > x_1 \wedge \bar{b} < x_2) = 2(1-x_1)2x_2$$

$$\begin{aligned} E_{\eta,\gamma}[A_1] &= 2x_1 - 1 + \frac{1}{2}\eta * 2(1-x_1)(1-2x_2) + \frac{1}{2} * 2(1-x_1)2x_2 \\ &= 2x_1 - 1 + \eta(1-x_1)(1-2x_2) + 2(1-x_1)x_2 \\ &= 2x_1 - 1 + \eta(1-x_1)(1-2x_2) + 2(1-x_1)x_2 \end{aligned}$$

Let us verify that $E_{\eta,\gamma}[A_1] \geq x_1$

$$2x_1 - 1 + \eta(1-x_1)(1-2x_2) + 2(1-x_1)x_2 \geq x_1$$

$$x_1 - 1 + \eta(1-x_1)(1-2x_2) + 2(1-x_1)x_2 \geq 0$$

$$(1-x_1)(-1 + \eta(1-2x_2) + 2x_2) \geq 0$$

$$\underbrace{(1-x_1)}_{\geq 0}\underbrace{(1-\eta)}_{\geq 0}\underbrace{(2x_2-1)}_{\leq 0} \leq 0$$

$$\Rightarrow E_{\eta,\gamma}[A_1] \leq x_1$$

By symmetry the same applies to x_2.

$$E_{\eta,\gamma}[A_2] = \frac{1}{2}(1+\eta) * \gamma(\bar{d} < x_1 \ \wedge \ \bar{b} < x_2) + \frac{1}{2} * \gamma(\bar{d} > x_1 \wedge \bar{b} < x_2)$$

$$E_{\eta,\gamma}[A_2] = \frac{1}{2}(1+\eta) * 2\left(x_1 - \frac{1}{2}\right)2x_2 + \frac{1}{2} * 2(1-x_1)2x_2$$

$$E_{\eta,\gamma}[A_2] = 2(1+\eta)\left(x_1 - \frac{1}{2}\right)x_2 + 2(1-x_1)x_2$$

$E_{\eta,\gamma}[A_2] \geq x_2$?

$$2(1+\eta)\left(x_1 - \frac{1}{2}\right)x_2 + 2(1-x_1)x_2 \geq x_2$$

$$\left(2(1+\eta)\left(x_1 - \frac{1}{2}\right) + 2(1-x_1)\right)x_2 \geq x_2$$

$$(+2\eta x_1 - \eta + 1)\, x_2 \geq x_2$$

$$\underbrace{(\eta \underbrace{(2x_1 - 1)}_{\geq 0} + 1)}_{\geq 1} x_2 \geq x_2$$

$$E_{\eta,\gamma}[A_2] \geq x_2$$

Now if we prove that $E_{\eta,\gamma}[A_1] \geq x_2$ and $E_{\eta,\gamma}[A_2] \leq x_1$, then by eq (3) $E_{\eta,\gamma}[A_1] \leq x_1$ and $E_{\eta,\gamma}[A_2] \geq x_2$ are sufficient condition to ensure that $x_{1,t} \geq x_{1,t+1}$ and $x_{2,t} \leq x_{2,t+1}$, in all points $\{(x_1, x_2) : x_1 > \frac{1}{2}, x_2 < \frac{1}{2}\}$.

Let us verify that $E_{\eta,\gamma}[A_1] \geq x_2$

$$2x_1 - 1 + \eta(1 - x_1)(1 - 2x_2) + 2(1 - x_1)x_2 > x_2$$

$$2x_1 - 1 + \eta(1 - x_1)(1 - 2x_2) + 2(1 - x_1)x_2 - x_2 > 0$$

$$2x_1 - 1 + (1 - x_1)(\eta(1 - 2x_2) + 2x_2) - x_2 > 0$$

$$2x_1 - 1 - x_2 + (1 - x_1)(\eta + 2x_2(1 - \eta)) > 0 \quad \textit{Always}$$

Let us verify that $E_{\eta,\gamma}[A_2] \leq x_1$

$$2(1 + \eta)\left(x_1 - \frac{1}{2}\right) x_2 + 2(1 - x_1)x_2 \leq x_1$$

$$2x_1x_2 - x_2 + 2\eta x_1x_2 - \eta x_2 + 2x_2 - 2x_1x_2 \leq x_1$$

$$x_2 + 2\eta x_1x_2 - \eta x_2 \leq x_1$$

$$x_2(1 - \eta + 2\eta x_1) \leq x_1$$

if $x_1 = 1$

$$x_2(1 + \eta) \leq 1 \quad \text{always } (\max\{x_2\} = 1/2, \max\{\eta\} = 1)$$

if $x_1 = \frac{1}{2}$

$$x_2 \leq \frac{1}{2} \quad \text{always}$$

Since $E_{\eta,\gamma}[A_1]$ has x_2 as lower bound and $E_{\eta,\gamma}[A_2]$ has x_1 as upper-bound, then in the region $x_1 \geq \frac{1}{2}, x_2 \leq \frac{1}{2}$ x_1 decrease and x_2 increase over time in all points $\{(x_1, x_2) : x_1 > \frac{1}{2}, x_2 < \frac{1}{2}\}$. The opposite holds in $\{(x_1, x_2) : x_1 < \frac{1}{2}, x_2 > \frac{1}{2}\}$, by symmetry. Notice that these two conditions ensure that the only possible equilibrium in these regions is $\left(\frac{1}{2}, \frac{1}{2}, \bar{f}, \bar{f}\right)$, which is a saddle.

□

Proof of Proposition 5

Again, to understand the dynamics we need to study $E_{\eta,\gamma}[A_1]$ and $E_{\eta,\gamma}[A_2]$ is the space $[0,1]^2$. In fact $E_{\eta,\gamma}[A_1]$ and $E_{\eta,\gamma}[A_2]$ affect directly the horizontal socialization dynamics (7) and the norms dynamics through (7) and (8). By symmetry we can focus only on regions $\{(x_1,x_2): x_1 > \frac{1}{2}, x_2 > \frac{1}{2}\}$ and $\{(x_1,x_2): x_1 > \frac{1}{2}, x_2 < \frac{1}{2}\}$.

- If $x_1 > \frac{1}{2}$ and $x_2 > \frac{1}{2}$

$$E_{\eta,\gamma}[A_1] = 1 * \gamma(\bar{b} < x_1) + \frac{1}{2}\eta * \gamma(\bar{b} > x_1 \wedge \bar{b} < x_2) + \frac{1}{2}\gamma * (\bar{b} > x_1 \wedge \bar{b} > x_2)$$

$$= 2x_1 - 1 + \frac{1}{2}\eta 4(1-x_1)(x_2 - \frac{1}{2}) + \frac{1}{2}4(1-x_1)(1-x_2)$$

Let us verify that $E_{\eta,\gamma}[A_1] \leq x_1$

$$2x_1 - 1 + 2\eta(1-x_1)\left(x_2 - \frac{1}{2}\right) + 2(1-x_1)(1-x_2) \leq x_1$$

$$x_1 - 1 + 2\eta(1-x_1)\left(x_2 - \frac{1}{2}\right) + 2(1-x_1)(1-x_2) \leq 0$$

$$(1-x_1)(2\eta\left(x_2 - \frac{1}{2}\right) + 2(1-x_2) - 1) \leq 0$$

$$(1-x_1)(2\eta x_2 - \eta + 1 - 2x_2) \leq 0$$

$$\underbrace{(1-x_1)}_{\geq 0}\underbrace{(1-2x_2)}_{\leq 0}\underbrace{(1-\eta)}_{\geq 0} \leq 0 \quad \text{Always}$$

By symmetry $E_{\eta,\gamma}[A_2] \leq x_2$, thus both x_1 and x_2 decrease overtime.

Since by eq (3) both $x_{1,t}$ and $x_{2,t}$ are convex combinations of $\theta_{1,t}$ and $\theta_{2,t}$ which directly depends on $E_{\eta,\gamma}[A_1]$ and $E_{\eta,\gamma}[A_2]$ respectively, then $E_{\eta,\gamma}[A_1] \leq x_1$ and $E_{\eta,\gamma}[A_2] \leq x_2$ are sufficient conditions to ensure that $x_{i,t} \geq x_{i,t+1}$ for all $i \in \{1,2\}$. By symmetry $x_{i,t} \leq x_{i,t+1}$ for all $i \in \{1,2\}$ when $x_1 \leq \frac{1}{2}$, $x_2 \leq \frac{1}{2}$. Thus, the only possible equilibrium in these regions is $\left(\frac{1}{2}, \frac{1}{2}, \bar{f}, \bar{f}\right)$, which is a saddle.

Let us assume that $x_1 > \frac{1}{2}$ and $x_2 < \frac{1}{2}$

$$E_{\eta,\gamma}[A_1] = 1 * \gamma(\bar{b} < x_1) + (1 - \frac{1}{2}\eta) * \gamma(\bar{b} > x_1 \ \wedge \ \bar{d} > x_2) + \frac{1}{2} * \gamma(\bar{b} > x_1 \wedge \bar{d} < x_2)$$

$$E_{\eta,\gamma}[A_1] = 2x_1 - 1 + 2(1 - \frac{1}{2}\eta)(1-x_1)(1-2x_2) + 2(1-x_1)x_2$$

Let us verify that $E_{\eta,\gamma}[A_1] > x_1$

$$2x_1 - 1 + 2(1 - \frac{1}{2}\eta)(1 - x_1)(1 - 2x_2) + 2(1 - x_1)x_2 > x_1$$

$$x_1 - 1 + 2(1 - \frac{1}{2}\eta)(1 - x_1)(1 - 2x_2) + 2(1 - x_1)x_2 > 0$$

$$(1 - x_1)(2(1 - \frac{1}{2}\eta)(1 - 2x_2) + 2x_2 - 1) > 0$$

$$(1 - x_1)(2 - \eta - 4x_2 + 2\eta x_2 + 2x_2 - 1) > 0$$

$$\underbrace{(1 - x_1)}_{\geq 0}\underbrace{(1 - 2x_2)}_{\geq 0}\underbrace{(1 - \eta)}_{\geq 0} \geq 0 \quad \text{Always}$$

$\Rightarrow E_{\eta,\gamma}[A_1] \geq x_1$ and by symmetry $E_{\eta,\gamma}[A_2] \leq x_2$

$$E_{\eta,\gamma}[A_2] = \frac{1}{2}(1 - \eta) * \gamma(\bar{b} < x_1 \ \wedge \ \bar{d} < x_2) + \frac{1}{2} * \gamma(\bar{b} > x_1 \wedge \bar{d} < x_2)$$

$$E_{\eta,\gamma}[A_2] = \frac{1}{2}(1 - \eta) * 2\left(x_1 - \frac{1}{2}\right) 2x_2 + \frac{1}{2} * 2(1 - x_1)2x_2$$

Let us verify that $E_{\eta,\gamma}[A_2] \leq x_2$

$$E_{\eta,\gamma}[A_2]] = \frac{1}{2}(1 - \eta) * 2\left(x_1 - \frac{1}{2}\right) 2x_2 + \frac{1}{2} * 2(1 - x_1)2x_2 \leq x_2$$

$$2(1 - \eta)\left(x_1 - \frac{1}{2}\right) x_2 + 2(1 - x_1)x_2 \leq x_2$$

$$2(1 - \eta)\left(x_1 - \frac{1}{2}\right) x_2 + 2(1 - x_1)x_2 \leq x_2$$

$$2x_1x_2 - x_2 - 2x_1x_2\eta + x_2\eta + 2x_2 - 2x_1x_2 \leq x_2$$

$$-2x_1x_2\eta + x_2\eta \leq 0$$

$$x_2\eta\underbrace{(1 - 2x_1)}_{\leq 0} \leq 0 \quad \textit{always}$$

We recall that

$$\begin{cases} x_{1,t+1} = \phi_{1,t+1}E_{\eta,\gamma}[A_{1,t}] + (1 - \phi_{1,t+1})E_{\eta,\gamma}[A_{2,t}] \\ x_{2,t+1} = \phi_{2,t+1}E_{\eta,\gamma}[A_{1,t}] + (1 - \phi_{2,t+1})E_{\eta,\gamma}[A_{2,t}] \end{cases}$$

Thus

$$\begin{cases} x_{1,t+1} \geq x_{1,t} & \iff \quad \phi_{1,t+1} E_{\eta,\gamma}[A^*_{1,t}] + (1-\phi_{1,t+1}) E_{\eta,\gamma}[A^*_{2,t}] \geq x_{1,t} \\ x_{2,t+1} \leq x_{2,t} & \iff \quad \phi_{2,t+1} E_{\eta,\gamma}[A^*_{1,t}] + (1-\phi_{2,t+1}) E_{\eta,\gamma}[A^*_{2,t}] \leq x_{2,t} \end{cases}$$

From which

$$\begin{cases} x_{1,t+1} \geq x_{1,t} & \iff \quad \phi_{1,t+1} \geq \frac{x_{1,t} - E_{\eta,\gamma}[A^*_{2,t}]}{E_{\eta,\gamma}[A^*_{1,t}] - E_{\eta,\gamma}[A^*_{2,t}]} \\ x_{2,t+1} \leq x_{2,t} & \iff \quad \phi_{2,t+1} \leq \frac{x_{2,t} - E_{\eta,\gamma}[A^*_{2,t}]}{E_{\eta,\gamma}[A^*_{1,t}] - E_{\eta,\gamma}[A^*_{2,t}]} \end{cases}$$

Since, in the space $\left[\frac{1}{2}, 1\right] \times \left[0, \frac{1}{2}\right]$, $E_{\eta,\gamma}[A_1] \geq x_1$ and $E_{\eta,\gamma}[A_2] \leq x_2$ for all t then as $t \to \infty$ $E_{\eta,\gamma}[A_1] \to 1$ and $E_{\eta,\gamma}[A_2] \to 0$ Thus

$$\begin{cases} x_{1,t+1} \geq x_{1,t} & \iff \quad \phi_{1,t+1} \geq x_{1,t} \\ x_{2,t+1} \leq x_{2,t} & \iff \quad \phi_{2,t+1} \leq x_{2,t} \end{cases}$$

Since $\phi_{1,t}$ and $\phi_{2,t}$ depends (among endogenous variables) only on $f_{1,t}$ and $f_{2,t}$ and both converges as $t \to \infty$ to $\bar{f}\eta$ and $\bar{f}(1-\eta)$ respectively, then $\phi_{1,t} \to \phi_1$ and $\phi_{2,t} \to \phi_2$. Thus in this region there are two possible equilibria

The opposite holds in $\{(x_1, x_2) : x_1 < \frac{1}{2}, x_2 > \frac{1}{2}\}$, by symmetry. □

□

Proof of Proposition 6

In specific regions of x, the dynamics of f is independent on the x and decoupled, namely the flexibility of one community does not depends on the one of the other. In particular it is

$$f_t = \begin{bmatrix} 1-\mu & 0 \\ 0 & 1-\mu \end{bmatrix} f_{t-1} + constant$$

Thus

$$\lim_{t\to\infty} f_t = f^*$$

Now we can consider the dynamics on x_t. For all possible f the dynamics is linear in x, the coefficient matrix depends on f^* and it is a stochastic matrix

$$x_t = (1-\lambda) \begin{bmatrix} p_{1,t}(f^*) & 1-p_{1,t}(f^*) \\ p_{2,t}(f^*) & 1-p_{2,t}(f^*) \end{bmatrix} x_{t-1} + constant$$

We can observe that the coefficient matrix remain stochastic for all time period, namely $\sum_j P_{i,j} = 1$. For example, moving one period ahead and looking at the first row of the coefficient matrix looking at the first row we have that

$$p_{1,t}p_{1,t-1} + (1 - p_{1,t})p_{2,t-1} + p_{1,t}(1 - p_{1,t-1}) + (1 - p_{1,t})(1 - p_{2,t-1})$$

$$p_{1,t}p_{1,t-1} + p_{2,t-1} - p_{1,t}p_{2,t-1} + p_{1,t} - p_{1,t}p_{1,t-1} + 1 - p_{2,t-1} - p_{1,t} + p_{1,t}p_{2,t-1} = 1$$

And thus as $t \to \infty$ we have

$$x_t - x^{*i} = (1 - \lambda)^t \mathbf{P}(t)(x_0 - x^{*i})$$

where $\mathbf{P}(t)$ is a stochastic matrix, and term $(1 - \lambda)^t$ converges to zero, therefore

$$\lim_{t \to \infty} x_t = x^*$$

And in all the region of the state space we converge to a steady state.
In the next section we characterize all the possible steady states of (1.10) and their possible basin of attraction.

□

Point Distribution: Steady States

In this section we discuss in general all the possible steady state with point distribution. Results hold in general.

We define e as a generic element of $\mathcal{E}$, thus we can enumerate the possible steady states as $e^1, e^2, ...$

Proposition 7 *[Convergence]*
Consider the norm and socialization level dynamics in (1.10). For all $\eta \in (0,1)$, $\bar{b} \in (0,1)$, $\bar{d} \in (0,1)$, $\bar{f} \in (0,1)$, $\lambda \in (0,1)$

- *$e^1 = (1, 1, \bar{f}, \bar{f})$ and $e^2 = (0, 0, \bar{f}, \bar{f}) \in \mathcal{E}$. The average actions at the steady state are, respectively, $E_\eta[A^*] = (1, 1)$ and $A^* = (0, 0)$.*
- *$e^3 = \left(\frac{1}{2}, \frac{1}{2}, \bar{f}, \bar{f}\right) \in \mathcal{E}$ if and only if the original 2x2 symmetric games has multiple Nash equilibria. The average actions at the steady state are $E_\eta[A^*] = \left(\frac{1}{2}, \frac{1}{2}\right)$*

Proposition 8 *[Polarization]*
Consider the norm and socialization level dynamics in (1.10). For all $\eta \in (0,1)$, $\bar{b} \in (0,1)$, $\bar{d} \in (0,1)$, $\bar{f} \in (0,1)$, $\lambda \in (0,1)$

- $e^4 = (\phi_1, \phi_2, \bar{f}\eta, \bar{f}(1-\eta)) \in \mathcal{E}$ *if and only if* $\phi \in R_{1_*,0_*}$. *The average actions at the steady state are* $E_\eta[A^*] = (1, 0)$.

- $e^5 = (1-\phi_1, 1-\phi_2, \bar{f}\eta, \bar{f}(1-\eta)) \in \mathcal{E}$ *if and only if* $1-\phi \in R_{0_*,1_*}$. *The average actions at the steady state are* $E_\eta[A^*] = (0, 1)$.

Proposition 9 *[Partial Convergence]*
Consider the norm and socialization level dynamics in (1.10). For all $\eta \in (0,1)$, $\bar{b} \in (0,1)$, $\bar{d} \in (0,1)$, $\bar{f} \in (0,1)$, $\lambda \in (0,1)$, *with* $\bar{b} < \bar{d}$

- $e^6_{\bar{d}} = \left(1 - \frac{1}{2}(1-\eta)(1-\phi_1), 1 - \frac{1}{2}(1-\phi_2)(1-\eta), \bar{f}\left(1 - \frac{1}{2}(1-\eta)^2\right), \bar{f}\left(1 - \frac{1}{2}\eta(1-\eta)\right)\right) \in \mathcal{E}$ *if and only if* $1 - \frac{1}{2}(1-\phi_2)(1-\eta) < \bar{d} < 1 - \frac{1}{2}(1 - \eta(1-\phi_1))$. *The average actions at the steady state are* $E_\eta[A^*] = \left(1, \frac{1}{2}(1+\eta)\right)$.

- $e^7_{\bar{d}} = \left(1 - \frac{1}{2}\eta\phi_1, 1 - \frac{1}{2}\eta\phi_2, \bar{f}\left(1 - \frac{1}{2}\eta(1-\eta)\right), \bar{f}\left(1 - \frac{1}{2}\eta^2\right)\right) \in \mathcal{E}$ *if and only if* $1 - \frac{1}{2}\eta\phi_1 < \bar{d} < 1 - \frac{1}{2}\eta\phi_2$. *The average actions at the steady state are* $E_\eta[A^*] = \left(1 - \frac{1}{2}\eta, 1\right)$.

- $e^8_{\bar{d}} = \left(\frac{1}{2}\eta\phi_1, \frac{1}{2}\eta\phi_2, \bar{f}\left(1 - \frac{1}{2}\eta(1-\eta)\right), \bar{f}\left(1 - \frac{1}{2}\eta^2\right)\right) \in \mathcal{E}$ *if and only if* $\frac{1}{2}\eta\phi_2 < \bar{b} < \frac{1}{2}\eta\phi_1$. *The average actions at the steady state are* $E_\eta[A^*] = \left(\frac{1}{2}\eta, 0\right)$.

- $e^9_{\bar{d}} = \left(\frac{1}{2}(1-\eta)(1-\phi_1), \frac{1}{2}(1-\eta)(1-\phi_2), \bar{f}\left(1 - \frac{1}{2}(1-\eta)^2\right), \bar{f}\left(1 - \frac{1}{2}\eta(1-\eta)\right)\right) \in \mathcal{E}$ *if and only if* $\frac{1}{2}(1-\eta)(1-\phi_1) < \bar{b} < \frac{1}{2}(1-\eta)(1-\phi_2)$. *The actions at the steady state are* $E_\eta[A^*] = \left(0, \frac{1}{2}(1-\eta)\right)$.

Proposition 10 *[Partial Polarization]*
Consider the norm and socialization level dynamics in (1.10). For all $\eta \in (0,1)$, $\bar{b} \in (0,1)$, $\bar{d} \in (0,1)$, $\bar{f} \in (0,1)$, $\lambda \in (0,1)$, *with* $\bar{d} < \bar{b}$

- $e^6_{\bar{b}} = \left(1 - \frac{1}{2}(1-\phi_1)(1+\eta), 1 - \frac{1}{2}(1-\phi_2)(1+\eta), \bar{f}(\frac{1}{2}(1+\eta^2)), \bar{f}(1 - \frac{1}{2}\eta(1+\eta))\right) \in \mathcal{E}$ *if and only if* $\bar{b} < 1 - \frac{1}{2}(1-\phi_1)(1+\eta) \wedge \bar{d} < 1 - \frac{1}{2}(1-\phi_2)(1+\eta)$. *The average actions at the steady state are* $E_\eta[A^*] = \left(1, \frac{1}{2}(1-\eta)\right)$.

- $e^7_{\bar{b}} = \left(1 - \phi_1(1 - \frac{1}{2}\eta), 1 - \phi_2(1 - \frac{1}{2}\eta), \bar{f}(1 - \frac{1}{2}\eta(1-\eta)), \bar{f}(\frac{1}{2}\eta(3-\eta))\right) \in \mathcal{E}$ *if and only if* $\bar{d} < 1 - \phi_1(1 - \frac{1}{2}\eta) \wedge \bar{b} < 1 - \phi_2(1 - \frac{1}{2}\eta)$. *The average actions at the steady state are* $E_\eta[A^*] = \left(\frac{1}{2}\eta, 1\right)$.

- $e^8_{\bar{b}} = \left(\phi_1(1 - \frac{1}{2}\eta), \phi_2(1 - \frac{1}{2}\eta), \bar{f}(1 - \frac{1}{2}\eta(1-\eta)), \bar{f}(\frac{1}{2}\eta(3-\eta))\right) \in \mathcal{E}$ *if and only if* $\phi_2(1 - \frac{1}{2}\eta) < \bar{d} < \phi_1(1 - \frac{1}{2}\eta) \wedge \bar{b} > \phi_1(1 - \frac{1}{2}\eta)$. *The average actions at the steady state are* $E_\eta[A^*] = \left(1 - \frac{1}{2}\eta, 0\right)$.

- $e^9_{\bar{b}} = \left(\frac{1}{2}(1+\eta)(1-\phi_1), \frac{1}{2}(1+\eta)(1-\phi_2), \bar{f}(\frac{1}{2}(1+\eta^2)), \bar{f}(1 - \frac{1}{2}\eta(1+\eta))\right) \in \mathcal{E}$ *if and only if* $\frac{1}{2}(1+\eta)(1-\phi_1) < \bar{d} < \frac{1}{2}(1+\eta)(1-\phi_2) \wedge \bar{b} > \frac{1}{2}(1+\eta)(1-\phi_2)$. *The average actions at the steady state are* $E_\eta[A^*] = \left(0, \frac{1}{2}(1+\eta)\right)$.

Corollary 10.1 *The steady state described in Proposition 8, 9, 10 and 11 are the only possible steady states of* (1.10). *Moreover*

- *If the game played is with complements,* $\bar{b} < \frac{1}{2} < \bar{d}$, (1.10) *has a minimum of three steady states,* (e^1, e^2, e^3)*, and a maximum of nine,* $(e^1, ..., e^9_{\bar{d}})$.
- *If the game played is with substitution,* $\bar{b} < \frac{1}{2} < \bar{d}$, (1.10) *has a minimum of three steady states,* (e^1, e^2, e^3)*, and a maximum of nine,* $(e^1, ..., e^9_{\bar{b}})$.
- *If the game played is a Prisoner Dilemma,* $\frac{1}{2} < \{\bar{b}, \bar{d}\}$, (1.10) *has a minimum of two steady states,* (e^1, e^2)*, and a maximum of six. The latter are* $(e^1, ..., e^5, e^7_{\bar{d}}, e^8_{\bar{d}})$ *if* $\bar{d} < \bar{b}$*, and* $(e^1, ..., e^5, e^6_{\bar{d}}, e^9_{\bar{d}})$ *if* $\bar{b} < \bar{d}$.

With point distribution $E_{\eta,\gamma}[A] = E_\eta[A]$, thus in order to prove the previous propositions we need to substitute $E_\eta[A]$ (listed in Figure 4) in equation (16) and (17).

Now given the definition of basin of attraction and proposition 8, 9, 10, 11 it is trivial to find all the possible basin of attraction. The possible steady states of (1.10) have the following basin of attraction:

1. $$B(e^1)\begin{cases} \ni R_{1_*,1_*} & \textit{always} \\ \ni R_{1_*,1} & \text{iff} \quad \bar{d} > \bar{b} \wedge e^6_{\bar{d}} \in R_{1_*,1_*} \\ \ni R_{1,1_*} & \text{iff} \quad \bar{d} > \bar{b} \wedge e^6_{\bar{d}} \in R_{1_*,1_*} \\ \ni R_{1_*,0_*} & \text{iff} \quad \bar{d} > \bar{b} \wedge e^4 \in R_{1_*,1} \wedge e^6_{\bar{d}} \in R_{1_*,1_*} \\ \ni R_{0_*,1_*} & \text{iff} \quad \bar{d} > \bar{b} \wedge e^5 \in R_{1,1_*} \wedge e^6_{\bar{d}} \in R_{1_*,1_*} \end{cases}$$

2. $$B(e^2)\begin{cases} \ni R & \text{iff} \quad e^2 \in R \\ \ni R_{1_*,0_*} & \text{iff} \quad e^4 \in R \\ \ni R_{0_*,1_*} & \text{iff} \quad e^5 \in R \\ \ni R_{1_*,1} & \text{iff} \quad \bar{d} > \bar{b} \wedge e^6_{\bar{d}} \in R \\ \ni R_{1,1_*} & \text{iff} \quad \bar{d} > \bar{b} \wedge e^7_{\bar{d}} \in R \\ \ni R_{0,0_*} & \text{iff} \quad \bar{d} > \bar{b} \wedge e^8_{\bar{d}} \in R \\ \ni R_{0_*,0} & \text{iff} \quad \bar{d} > \bar{b} \wedge e^9_{\bar{d}} \in R \\ \ni R_{1_*,0} & \text{iff} \quad \bar{b} > \bar{d} \wedge e^6_{\bar{b}} \in R \\ \ni R_{0,1_*} & \text{iff} \quad \bar{b} > \bar{d} \wedge e^7_{\bar{b}} \in R \\ \ni R_{1,0_*} & \text{iff} \quad \bar{b} > \bar{d} \wedge e^8_{\bar{b}} \in R \\ \ni R_{0_*,1} & \text{iff} \quad \bar{b} > \bar{d} \wedge e^9_{\bar{b}} \in R \end{cases}$$

3. $$B(e^3)\begin{cases} \ni R_{0_*,0_*} & \mathit{always} \\ \ni R_{0,0_*} & \text{iff} \quad \bar{d} > \bar{b} \wedge e^8_{\bar{d}} \in R_{0_*,0_*} \\ \ni R_{0_*,0} & \text{iff} \quad \bar{d} > \bar{b} \wedge e^9_{\bar{d}} \in R_{0_*,0_*} \\ \ni R_{1_*,0_*} & \text{iff} \quad \bar{d} > \bar{b} \wedge e^4 \in R_{0,0_*} \wedge e^8_{\bar{d}} \in R_{0_*,0_*} \\ \ni R_{0_*,1_*} & \text{iff} \quad \bar{d} > \bar{b} \wedge e^5 \in R_{0_*,0} \wedge e^9_{\bar{d}} \in R_{0_*,0_*} \end{cases}$$

4. $$B(e^4)\begin{cases} \ni R_{1_*,0_*} & \text{iff} \quad e^4 \in E \\ \ni R_{1_*,0} & \text{iff} \quad \bar{b} > \bar{d} \wedge e^6_{\bar{b}} \in R_{1_*,0_*} \\ \ni R_{1,0_*} & \text{iff} \quad \bar{b} > \bar{d} \wedge e^8_{\bar{b}} \in R_{1_*,0_*} \end{cases}$$

5. $$B(e^5)\begin{cases} \ni R_{0_*,1_*} & \text{iff} \quad e^5 \in E \\ \ni R_{0,1_*} & \text{iff} \quad \bar{b} > \bar{d} \wedge e^7_{\bar{b}} \in R_{0_*,1_*} \\ \ni R_{0_*,1} & \text{iff} \quad \bar{b} > \bar{d} \wedge e^9_{\bar{b}} \in R_{0_*,1_*} \end{cases}$$

6. $$B(e^6_{\bar{d}})\begin{cases} \ni R_{1_*,1} & \text{iff} \quad \bar{d} > \bar{b} \wedge e^6_{\bar{d}} \in E \\ \ni R_{1_*,0_*} & \text{iff} \quad \bar{d} > \bar{b} \wedge e^4 \in R_{1_*,1} \end{cases}$$

7. $$B(e^7_{\bar{d}})\begin{cases} \ni R_{1,1_*} & \text{iff} \quad \bar{d} > \bar{b} \wedge e^7_{\bar{d}} \in E \\ \ni R_{0_*,1_*} & \text{iff} \quad \bar{d} > \bar{b} \wedge e^5 \in R_{1,1_*} \end{cases}$$

8. $$B(e^8_{\bar{d}})\begin{cases} \ni R_{0,0_*} & \text{iff} \quad \bar{d} > \bar{b} \wedge e^8_{\bar{d}} \in E \\ \ni R_{1_*,0_*} & \text{iff} \quad \bar{d} > \bar{b} \wedge e^4 \in R_{0,0_*} \end{cases}$$

9. $B(e^9_{\bar{d}}) \begin{cases} \ni R_{0_*,0} & \text{iff} \quad \bar{d} > \bar{b} \wedge e^9_{\bar{d}} \in E \\ \ni R_{0_*,1_*} & \text{iff} \quad \bar{d} > \bar{b} \wedge e^5 \in R_{0_*,0} \end{cases}$

10. $B(e^6_{\bar{b}}) \begin{cases} \ni R_{1_*,0} & \text{iff} \quad \bar{b} > \bar{d} \wedge e^6_{\bar{b}} \in E \\ \ni R_{1_*,0_*} & \text{iff} \quad \bar{b} > \bar{d} \wedge e^4 \in R_{1_*,0} \end{cases}$

11. $B(e^7_{\bar{b}}) \begin{cases} \ni R_{0,1_*} & \text{iff} \quad \bar{b} > \bar{d} \wedge e^7_{\bar{b}} \in E \\ \ni R_{0_*,1_*} & \text{iff} \quad \bar{b} > \bar{d} \wedge e^5 \in R_{0,1_*} \end{cases}$

12. $B(e^8_{\bar{b}}) \begin{cases} \ni R_{1,0_*} & \text{iff} \quad \bar{b} > \bar{d} \wedge e^8_{\bar{b}} \in E \\ \ni R_{1_*,0_*} & \text{iff} \quad \bar{b} > \bar{d} \wedge e^4 \in R_{1,0_*} \end{cases}$

13. $B(e^9_{\bar{b}}) \begin{cases} \ni R_{0_*,1} & \text{iff} \quad \bar{b} > \bar{d} \wedge e^9_{\bar{b}} \in E \\ \ni R_{0_*,1_*} & \text{iff} \quad \bar{b} > \bar{d} \wedge e^5 \in R_{0_*,1} \end{cases}$

Now we can consider all the possible equilibria e basin of attraction and we can easily check that

$$\bigcup_{\forall e^i \in E} B(e^i) = [0,1]^2 \times [0,1]^2$$

Thus, independently on parameters $\bar{b}, \bar{d}, \eta, \bar{f}$ the dynamics (1.10) always converge to a stable steady state, no matter the starting point.

□

Chapter 2

Cultural Transmission with Incomplete Information: Parental Self-Efficacy and Group Misrepresentation[1]

Abstract

This paper introduces incomplete information in the standard model of cultural transmission (Bisin and Verdier, 2001). We show that if parents are not fully aware of own group size and about the efficiency of the cultural transmission technology, they may end up to sustain wrong conjectures about both quantities. This translates in having complementarity instead of substitution between optimal socialization effort and own population share. In the long run, if conjectures are shaped by cultural leaders who want to maximize the presence of the trait of their own community in the next period, conjectures are characterized by a negative bias so that agents tend to underrepresent own group in the population. Our main finding is that, depending on the magnitude of the bias, the dynamics can display stable or unstable polymorphic equilibria, or just a stable homomorphic equilibrium.

Journal of Economic Literature Classification Numbers: A14, C72, J15, D10, Z10, Z13
Keywords: Cultural Transmission, Group Under-Representation, Self-Confirming Equilibrium, Incomplete Information, Cultural leaders

[1]This chapter is joint with Fabrizio Panebianco, Università Cattolica del Sacro Cuore. E-mail: fabrizio.panebianco@unicatt.it

2.1 Introduction

It is a well-established fact that members of cultural or ethnic groups find it hard to have an unbiased perception of own population share in the society, even if the correct information is publicly available. In 2016 the magazine "The Economist" analyzed the perception that European citizens of different countries have about the share of muslim population in own country[2]. It turned out that this perception is extremely biased. For example, the less biased ones seem to be the Germans who think muslims to be 19% of the population, while they are just about 6%. The most biased are Hungarians, who think muslims to be 7% while they are 70 times less, 0.1%. A recent paper by Alesina et al. (2018) finds evidence that citizens strongly overestimate the share of migrants in the population. In some cases, as for the US, while migrants accounts for a 10% of the population, people think they are almost 40%. This strong misperception, paired with the *populist* story of immigrants' invasion, may have consequences on the way agents of different groups decide, for example, to transmit own cultural values to other members, and, when it comes to cultural traits, can affect the long run composition of the society.

Considering the issue of values' transmission, it is well know in the social psychological literature that the efficacy of parental transmission, that is the technology by which parents try to transmit own traits and opinions to offspring, may not be perfect or frictionless. The true parental efficacy is difficult to test since it depends on both external exogenous factors, such as neighborhood composition, child characteristics, or other ecological variables (Belsky et al., 1984), and on internal endogenous characteristics, such as parental beliefs about their own efficacy. The latter is known as "Parental Self-Efficacy" (PSE), which can be defined as people's beliefs about their capabilities to organize and execute a set of tasks related to parenting a child. High levels of "Parental Self-Efficacy" are associated with higher quality of parent-child interactions that leads to better offsprings' development, higher parental involvement and efficacy and lower stress (Bandura, 1993; Coleman and Karraker, 2000; de Montigny and Lacharité, 2005, among others).

We take into consideration the possibile misperception of own group size, and Parental Self-Efficacy issues, by introducing incomplete information in the standard model of cultural transmission (Bisin and Verdier, 2001). We consider a cultural transmission model in which parents of two different cultural groups exert some effort to try to induce own type to children. However, differently from standard literature, we introduce incomplete information in two different ways: we assume that parents ignore the actual population shares, and they also

[2] *The Economist*, "Islam in Europe: perception and reality", 23/03/2016

ignore how efficiently they are able to send messages to children during the socialization process.This double uncertainty is motivated by the fact that, on one hand, can be very costly to check the exact population share in the neighborhood where the children interact; on the other hand, parental true effectiveness is by definition unobservable and parents can only try to infer it looking at offspring's outcome.[3] Parents just have conjectures about both, and then exert socialization efforts maximizing a subjective expected utility. These two elements are both new with respect to the literature in which it is usually assumed that parents know population shares and the transmission technology is common knowledge and normalized to a zero-frictions technology. We show that if parents are not fully aware of own group size and about the efficiency of the cultural transmission technology, they may end up to sustain wrong conjectures about both quantities. This translates in long run consequences about composition of the society that may reverse the standard predictions of the model.

Socialization is a complex process that makes parents very attentive about their choices. Moreover, since parents may ignore their group population shares and their own efficacy in parental transmission, an implicit process of learning occurs during parenting to try to have correct conjectures about these. To take this issues into account, we consider parents that try not to have wrong conjectures about the true parameters, to efficiently transmit their values.In details, during the socialization process parents receive some form of feedback from children about how much they have been overall convinced by each trait during the socialization process. This feedback enables parents to make some inference about own conjectures. Only conjectures that are compatible with the feedback received can be sustained. We model this with an equilibrium concept that fits very well with this situation, the self-confirming (or conjectural) equilibrium (Battigalli and Guaitoli, 1988; Fudenberg and Levine, 1993a). Indeed, this equilibrium requires that, under incomplete information, agents maximize their subjective expected utility and have conjectures that must be confirmed by the feedback they receive, but they may be wrong. Although this equilibrium concept is static, as we discuss in the paper, this equilibrium concept has also a strong learning foundation that requires repeated parent-child interaction where parents revise their conjectures until the parental experience does not confirm them.[4] We strongly believe that it is a natural way to model cultural transmission processes which are characterized by repeated parent-child interaction and parents' learning process.

The first set of results we have regards the static setting, that is the choice of a single gen-

[3]Also in social psychological literature, there are no data nor measures about the true parental effectiveness, but only about PSE.

[4]In details, any form of adaptive learning process, if converges, it does so to a self-confirming equilibrium (Milgrom and Roberts, 1991).

eration. At first, in Section 3, we characterize the set of self-confirming equilibria, showing that a higher "Parental Self-Efficacy" induces higher socialization effort, as documented by social psychological literature. Then we discuss how the key concepts of cultural substitution and cultural complementarities change in our setting. While with complete information there is no difference between conjectures and true parameters, with incomplete information this difference is crucial. We propose the definition of "conjectured cultural substitution (complementarity)" - optimal socialization efforts are decreasing (increasing) in the conjecture about own population shares - as opposed to the "actual cultural substitution (complementarity)" - optimal socialization efforts are decreasing (increasing) in own population shares. Our main finding is that it is possible to obtain "actual cultural complementarity" in the Bisin and Verdier (2001) cultural transmission mechanism if incomplete information is considered. Moreover, whenever there is "conjectured cultural substitution (complementarity)" then there is "actual cultural complementarity (substitution)". In particular, if agents underestimate the share of their traits in the society there is cultural substitution with respect the conjecture and, surprisingly, cultural complementarity with respect the true population share. On the other hand, whenever agents overestimate their own group in the society we obtain multiple equilibria and, depending on the choice, both cultural complementarity and substitution can exists. We also show that no agent can largely overestimate own group size and still have own conjectured confirmed, while underestimation of any magnitude can be sustained and confirmed. We study the welfare loss associated with incomplete information with respect to the standard case in Section 5. We show that the loss is increasing in the difference between the conjecture and the true parameter. We also discuss how intolerance relates with welfare loss.

We then move to the analysis of the dynamic in section 6, introducing a minimal model of leadership. We postulate the existence of two cultural leaders, one for each cultural group.[5] Leaders can choose to induce a positive or negative bias in the conjectures of their community members. Each cultural leader, in order to maximize the share of the trait of her own community in the next period, always choose to instill a negative bias in agents belonging to their cultural group. The implications are particularly rich. We identify thresholds for the magnitude of biases of the two communities. Depending on the biases' magnitude with respect the threshold values, all possible social compositions can arise in the long run. In particular, if the bias of one cultural group overcome the corresponding threshold, then there exist a stable equilibrium in which the whole society converge to that culture. The most important result, with respect previous literature, is that if the bias of both groups do not overcome the threshold value then, despite having actual cultural complementarity in socialization choice there exists

[5]In recent years the study of cultural leaders has become of interest in the cultural transmission literature, for example Verdier and Zenou (2015), Prummer and Siedlarek (2017) and Verdier and Zenou (2018)

a stable equilibrium with cultural heterogeneity. In the opposite case an unstable polymorphic equilibrium is shown and long run cultural homogeneity is the result. This is particularly relevant because, in previous cultural transmission papers (Bisin and Verdier, 2001; Cheung and Wu, 2018), cultural complementarity leads always to cultural homogeneity. We further discuss the role of different magnitude of population conjectures' bias showing that an higher bias translate in a higher population share in the long-run. Moreover, the presence of one group in the society at the steady state positively depends on the intolerance and effectiveness as parents of agents belonging to that specific group.

2.2 The Model

Consider a society composed of a continuum of agents. Each agent belongs to one cultural group. Each group is characterized by a specific discrete cultural trait. Let the set of traits be $I := \{a, b\}$ and the fraction of individuals with trait $i \in I$ be q^i. We assume that in each group all agents are equal. Then, with a little abuse of notation we refer to i as the representative agent displaying trait $i \in I$. In each period, each agent reproduces asexually giving birth to just one child. Children are born without any specific trait, and traits are acquired during the cultural transmission process. Cultural transmission from one generation to the next one is a probabilistic process that is the result of a *vertical socialization* step, that transmits parental trait to child with probability d^i, and an *oblique socialization* step, that transmits with probability $1 - d^i$ a random trait of the population.[6] Parents exert an effort $\tau^i \in \mathcal{T} := [0, 1]$ to induce own type to child in the vertical socialization step. Then $d^i := \varphi(\tau^i)$ where $\varphi : \mathcal{T} \to [0, 1]$ and is increasing in τ^i.

Parental Efficacy In the standard cultural transmission literature, from Bisin and Verdier (2001) on, it is assumed that parental effort translates without friction to *vertical socialization*, namely $\varphi(\tau^i) = \tau^i$. However, the efficacy of parental transmission may not be perfect due to several factors, such as neighborhood composition, child characteristics, or other ecological variables (Belsky et al., 1984). For this reason, we consider a generalization of $\varphi(\tau^i)$ allowing for an efficacy parameter $\alpha^i \in \mathbb{R}_+$, defining $d^i := \varphi(\tau^i) = \alpha^i \tau^i$. The parameter α^i can reduce the efficacy of effort in case of frictions, but in some cases the vertical socialization technology can be such that it magnifies the effects of the parental effort.[7] To simplify the analysis, we assume the efficiency to be group specific. Call $S := [0, 1]^2$ and, for each $i \in I$, let $s^i \in S$ be a generic pair (α^i, q^i). We define the following consequence function mapping from the triple

[6] We remand to Cavalli-Sforza and Feldman (1981) and Bisin and Verdier (2001) for the terminology.

[7] Even if, with a generic $\alpha^i \in \mathbb{R}_+$, we may have that $\alpha^i \tau^i > 1$, we will show later that it will never be the case so that $\varphi : \mathcal{T} \to [0, 1]$ is always true.

of effort, efficiency, and population shares to transition probability p^{ii}, that is the probability that a given parent i gets a child of the same type. For each $i \in I$,

$$\begin{array}{rccl} p^{ii}: & \mathcal{T} \times S & \rightarrow & [0,1] \\ & (\tau^i, \alpha^i, q^i) & \mapsto & \alpha^i \tau^i + (1-\alpha^i\tau^i) q^i \end{array} \tag{2.1}$$

We also define $p^{ij} := 1 - p^{ii}$ that is the probability that own child is socialized to the different trait. This is given by $p^{ij} = (1-\alpha^i\tau^i)(1-q^i)$.

Incomplete information Differently from standard literature in cultural transmission, we assume that parents have *incomplete information* about the efficiency of their effort, α^i, and about the composition of the social environment the child is embedded into, q^i.[8] Indeed parents may not be sure about how much the messages they send to children during the vertical socialization are effective, and they may also ignore the exact composition of the society in terms of groups shares. Then, each parent has conjectures about s^i and, given these conjectures, produces a *subjectively optimal* effort to try to induce own type to the child. For each $i \in I$, define as $\hat{s}^i := (\hat{\alpha}^i, \hat{q}^i) \in S$ the pair of conjectured effort efficiency and population shares.[9] Given $i \in I$, and given $\hat{s}^i$, conjectures induce a *conjectured transition probability* function, describing what parent i thinks is the probability that own child is socialized to the same trait as her own.

$$\begin{array}{rccl} \hat{p}^{ii}: & \mathcal{T} \times S & \rightarrow & [0,1] \\ & (\tau^i, \hat{\alpha}^i, \hat{q}^i) & \mapsto & \hat{\alpha}^i \tau^i + (1-\hat{\alpha}^i\tau^i) \hat{q}^i \end{array} \tag{2.2}$$

As we did for transition probabilities, we define $\hat{p}^{ij} := 1 - \hat{p}^{ii} = (1-\hat{\alpha}^i\tau^i)(1-\hat{q}^i)$.

Notice that parents are not able to observe their own actual efficacy, α^i, but they can have information only about the effect of their action on offspring outcome (p^{ii} in this model). Thus,

[8]Note that, although children may have a neighborhood composition different from the overall population shares due to homophily, in this framework we choose to ignore this issue. Even if homophily may be somehow relevant, the main contribution of this paper regards how parental cultural transmission works when oblique socialization effects, however defined, are not known to the parents, together with parental effort efficacy. In the first part of the analysis, when no population dynamics occurs, it would be indifferent to allow for homophily or not, since what is relevant is the overall effect of oblique socialization on transition probabilities. When it comes to the population dynamic, the dynamics with and without homophily would be a bit different. However, we decided to ignore this issue since the presence of homophily does not change the quality of the results, but makes the framework fr too complicated distracting from the main research question.

[9]Note that we consider here just deterministic conjectures. We can potentially have probabilistic conjectures of the form $\hat{s}^i \in \Delta(S)$. However, given the structure of the utility function we describe below, for every probabilistic conjecture there exists one deterministic conjecture inducing the same subjectively optimal action. Then, without loss of generality, we decided to focus on deterministic conjectures.

they form conjectures about $\hat{\alpha}^i$.[10] $\hat{\alpha}^i$ is what in the social psychological literature is called "Parental Self-Efficacy", namely people's beliefs about their capabilities to organize and execute a set of tasks related to parenting a child (Bandura, 1993; Coleman and Karraker, 2000; de Montigny and Lacharité, 2005, among others).

Subjective expected utility maximization We assume that each parent $i \in I$ prefers having a child of own type than one of a different type. As standard, for every $i \in I$ and $j \in I \setminus \{i\}$, we model these preferences as a vector $(V^{ii}, V^{ij}) \in [0,1]^2$ where $V^{ii} > V^{ij}$. For each $i \in I$, let $\Delta V^i := V^{ii} - V^{ij}$.

Parents choose the level of socialization effort using *imperfect empathy*, namely they evaluate the types of children using their own preferences (Bisin and Verdier, 2001). Assuming quadratic socialization costs, parents maximize their subjective expected utility given own conjectures. Then, for every $i \in I$, and $j \in I \setminus \{i\}$ we get the following problem

$$\max_{\tau^i \in [0,1]} \mathbb{E}^i_{\hat{p}^{ii}}[u^i] = \hat{p}^{ii} V^{ii} + \hat{p}^{ij} V^{ij} - \frac{1}{2}(\tau^i)^2 \tag{2.3}$$

Confirmed conjectures Since parents try to do their best to successfully socialize own children, we assume that they try not to have wrong conjectures. The best they can do is not to have conjectures that contrast with some piece of evidence they can get. To do so we assume that each parent $i \in I$, during the socialization process, receives some message m^i from own child, and that, in equilibrium, she must have conjectures that are not in contrast with the message received. For simplicity we assume that the message received by parents is exactly the transition probability, that is $m^i = p^{ii}$. The assumption here is that parents are able to perfectly observe the true p^{ii}. We thus assume that parents, during the whole socialization process, are able to understand through deep communication with own children how much the children are convinced by parental traits in a $[0,1]$ scale.[11] This is, we think, a particularly relevant aspect of socialization, somehow ignored in standard literature, since parents are not blind and do not just discover the type of child at the end of the process, but interact with children during the socialization and may get some messages about how children are prone to get one trait against the other one.[12] Therefore, the information obtained by the parent $i \in I$ can be described by a feedback function $f : \mathcal{T} \times S \to [0,1]$. If a parent $i \in I$ receives a particular message m^i, what i

[10] Also in social psychology literature does not exists a measure of objective efficacy. In particular,

[11] Coleman and Karraker (2000)) outlines several possible mechanisms through which parental self-efficacy $\hat{\alpha}$ develops, in our model we focus on the actual parental experiences, in fact, the feedback provided from adult-child interactions is considered a key determinant of the formation of parental self-efficacy.

[12] Below in the paper we provide an interpretation of the proposed concept of equilibrium and about confirmed conjectures in which parents update time by time their effort depending on the message sent by children, and converge to the equilibrium effort For the time being we assume that the effort is chosen by each parent once and for all at the beginning of each period.

can infer, conditioned on her socialization effort τ^i, is that the set of conjectures consistent with the message is given by $f^{-1}_{i,\tau^i}(m^i) := \{(\alpha^i, q^i) : f_i(\tau^i, \alpha^i, q^i) = m^i\}$. This is the set of conjectures compatible with the message received. Notably, the true parameters must be part of this set, but this set will generically be larger.

Selfconfirming equilibrium We now define our equilibrium concept for this problem. We use the notion of selfconfirming equilibrium in which parents produce a subjectively optimal socialization effort, and the conjectures on which maximization is based are confirmed, that is, conjectures are compatible with the message received.[13] Let $r : S \to \mathbb{R}_+$ be the best response operator, so that $r(\hat{\alpha}^i, \hat{q}^i)$ is the subjectively optimal effort agent $i \in I$ exerts if she displays conjectures $(\hat{\alpha}^i, \hat{q}^i)$.[14]

Definition A profile $(\tau^i, \hat{\alpha}^i, \hat{q}^i)_{i\in I}$ of socialization choice and conjectures is a selfconfirming equilibrium at $(\alpha_i, q_i, \Delta V^i)_{i\in I}$ if, for each $i \in I$

1. (subjective rationality) $\tau^i \in r(\hat{\alpha}^i, \hat{q}^i)$
2. (confirmed conjectures) $(\hat{\alpha}^i, \hat{q}^i) \in f^{-1}_{i,\tau^i}(m^i)$

The first condition simply states that agents exert an effort that is an optimal response given their conjectures. The second condition implies that these conjectures must be confirmed by the message received. That is to say that if parents have these conjectures, then even if they are wrong, they cannot infer that they are wrong from the message received from the child. The main intuition behind this equilibrium concept is that if parents have incomplete information they can just act depending on some conjectures they have. Nothing ensures that these conjectures are correct. However, since parents get some feedback, they just keep conjectures that are compatible with the signals received. However, there the signal may be compatible with multiple conjectures, and some of them may be wrong.

As anticipated before, α^i is empirically impossible to observe given that parents may just observe the effect of their socialization effort and, as we assumed here, how much children are convinced about the different traits. By no way parent may have a direct and objective measure of their efficacy. Thus parents use received signals about population shares and transition probabilities to form conjectures about their self-efficacy. We have assumed that talking with their children, parents are perfectly able to deduce the correct transition probability p^{ii}. In such a case, as we show in the next section, parents manage to have correct conjectures about

[13] Battigalli and Guaitoli (1988), Fudenberg and Levine (1993a), Battigalli et al. (2015), Battigalli (2018).

[14] At this stage we have not proved that the correspondence r is single-valued. However, given the concavity of the problem this is ensured and we define directly from here $r : S \to [0, 1]$ as a function and not as a correspondence.

self-efficacy if and only if the signal about population share is correct. If we relax this assumption and messages on the transition probability are noisy, similar results occur, parents end up to correct conjectures $\hat{\alpha}^i$ if and only if signals about population share are noiseless.

Learning foundation of Selfconfirming equilibrium The selfconfirming equilibrium has a learning foundation that makes it particularly fit to model parental socialization choices. A selfconfirming equilibrium can be seen as the steady state of an adaptive learning dynamics in which conjectures are updated given the feedback agents receive at each period, when agents maximize their instantaneous expected utility (Fudenberg and Levine, 1993b; Milgrom and Roberts, 1991; Battigalli et al., 1992). In terms of our socialization process it is as if the socialization time is composed of infinitely many periods, and the child adopts a type at the end of all the periods. At the beginning of each period, each parent fixes her socialization choice. Then the vertical and oblique socialization schemes proceed as previously described, with children producing messages for parents. Each message is how much the child has been convinced, in that period, by each trait. Parents observe the message and update their conjectures, with an arbitrary adaptive learning dynamics. Then the process restarts again with a new round of vertical and oblique socialization. A fixed point of this adaptive learning process is a selfconfirming equilibria, where agents end up to play actions that are sustained by conjecture confirmed by experience. Thus if this process converges, then it must converge to a selfconfirming equilibrium. This mimics the fact that parents during socialization actually change their efforts depending on the feedback they can get from children. We highlight the fact that this learning foundation is just an interpretation of the selfconfirming equilibrium that is relevant for its introduction in a cultural transmission model with incomplete information. However the model we present here is not a learning model but, in each cohort, it is a static model that uses this equilibrium concept.

Notice that, as shown by Milgrom and Roberts (1991), the convergence to selfconfirming equilibrium does not depend on specific learning rules but holds for any possible adaptive learning. A process is consistent with adaptive learning if the players can find a way to justify choices in terms of the past realizations. Best Reply Dynamics, Fictitious Play, and Bayesian learning are all example of adaptive learning.[15] Therefore, selfconfirming equilibrium, although static, is consistent with all possible parental learning processes of this kind and it is the best way to capture the implicit process of learning that occurs during parenting.

[15]It is, thus, important to underline how selfconfirming equilibrium is the fixed point of the learning process even if agents update their conjecture as a Bayesian. This is relevant because it show that even fully rational Bayesian parents may have, in equilibrium, wrong conjecture about their efficacy and population shares. Intuitively, given parents' conjectures optimal effort levels can induce outcome that confirm the previous conjectures, in the case, parents do not revise their conjectures (for a deeper discussion see for example Battigalli, 2018).

2.3 Characterization of Equilibra

We can now characterize the set of selfconfirming equilibria.[16]

Proposition 1 *A selfconfirming equilibrium at* $(\alpha^i, q^i, \Delta V^i)_{i \in I}$ *of the socialization game is a profile* $(\tau^i, \hat{\alpha}^i, \hat{q}^i)_{i \in I}$ *in which, for each* $i \in I$

$$\tau^i = \hat{\alpha}^i (1 - \hat{q}^i) \Delta V^i \tag{2.4}$$

$$(\hat{\alpha}^i, \hat{q}^i) : \qquad \hat{\alpha}^i (1 - \hat{q}^i)[\alpha(1 - q^i) - \hat{\alpha}(1 - \hat{q}^i)] \Delta V^i = \hat{q}^i - q^i \tag{2.5}$$

Proof in the Appendix. □

The socialization effort positively depends on the "Parental Self Efficacy" $\hat{\alpha}^i$. This result is in line with the social psychological literature where, as outlined in Coleman and Karraker (2000), high parenting self-efficacy has been found to predict higher parental effort and performance.[17]

Notice that even if parents correctly observe p^{ii} from the messages they receive, this is not enough to make a correct inference about the underlying parameters (α^i, q^i). In fact, since the feedback function f is surjective, agents are not able to derive the exact parameter values but only the locus of points (τ^i, α^i) consistent of the feedback. We can observe that if an agent has a correct conjecture about the ratio of own trait in the population, namely $\hat{q}^i - q^i = 0$, then the only locus of points that supports *selfconfirming* equilibrium is $(\hat{\alpha}^i, \hat{q}^i) = (\alpha^i, q^i)$, namely parent has correct conjectures and the socialization choice described in (2.3) boils down to the standard (Bisin and Verdier, 2001) result

$$\tau^i_{bv} = \alpha^i (1 - q^i) \Delta V^i \tag{2.6}$$

Namely the Bisin and Verdier (2001) socialization choice, τ^i_{bv}, is nothing but the selfconfirming socialization choice when agents have correct conjectures about effectiveness and trait's share in the population. From now on we will use τ^i_{bv} as a benchmark for our analysis. We now go further in the equilibrium characterization analyzing how conjectures are shaped and the derived selfconfirming efforts differ from the standard literature ones.

[16] Notice that we define it as a *socialization game* even if there is no interaction among parents.

[17] For example, responsivity to children's needs (Donovan and Leavitt, 1985; Donovan et al., 1997; Unger and Wandersman, 1985), engagement in direct parenting interactions (Mash and Johnston, 1983), and active parental coping orientations (Wells-Parker et al., 1990).

We can interpret our model as if parents receive a bidimensional signal, over p^{ii} and q^i, where the first one is always correct. Different signals lead to different (confirmed) conjectures $\hat{\alpha}^i$, which is correct whenever the signal about q^i is correct. Notice that if parents do receive any signals about q^i, then $\hat{q}^i$ is nothing but an exogenous conjecture and together with $\hat{\alpha}^i$ should satisfy conditions of Proposition 1. Assuming absurdly that parents are perfectly able to observe their own self-efficacy, and thus conjectures $\hat{\alpha}^i$ are correct, (as in Bisin and Verdier (2001)), then $\hat{q}^i = q^i$ whenever the signal about $\hat{p}^{ii}$ is correct. In such a case parents may have incorrect conjectures about population share if only if they are not able to have correct guesses of transition probabilities of offsprings.

Let us start with the analysis of the conjectures in equilibrium.
Figure 2.1 provides a generic representation of (2.5), describing how equilibrium conjectures are shaped. In what follows we decide to characterize conjectures $\hat{\alpha}^i$, given $\hat{q}^i$. Indeed, fixing a conjecture $\hat{q}^i$, we study which $\hat{\alpha}^i$ satisfies (2.5). First of all if $\hat{q}^i$ is too large with respect to q^i

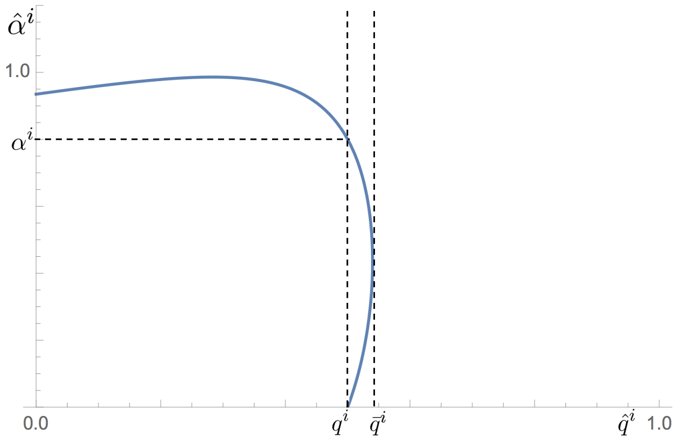


Figure 2.1: $\alpha^i = 0.8$, $\Delta V^i = 1$, $q^i = 0.4$

there are no conjectures compatible with the confirmation requirement. This is to say that an agent, in equilibrium, cannot think, and be confirmed in her conjecture, her group to be too large with respect to what it is in reality. In details, defining $\bar{q}^i := q^i + \frac{\tau_{bv}^{i2}}{4\Delta V^i}$, then if $\hat{q}^i > \bar{q}^i$ no confirmed conjectures can be find. Interestingly notice that, on the contrary, there is no lower bound to have confirmed conjectures, so that agents may have arbitrarily low conjectures on $\hat{q}^i$ and still find a $\hat{\alpha}^i$ confirming them. In other words, if agents strongly over-represent own group, they may find they are committing a mistake, while if they under-represent it, and maybe strongly under-represent it, they may not have a way to understand they are wrong. Moreover, if $\hat{q}^i \in [q^i, \bar{q}^i]$ two possible $\hat{\alpha}^i$ satisfy the confirmed conjectures requirement. If $\hat{q}^i < q^i$ then only a single $\hat{\alpha}^i$ satisfies the requirement.

We now formalize the discussion characterizing the Parental Self-Efficacy $\hat{\alpha}$, namely the conjecture on α, that can sustain selfconfiming equilibrium. Consider equation in (2.5) expressing the condition for confirmed conjectures. Let $\xi^i := \sqrt{4(q^i - \hat{q}^i)\Delta V^i + \alpha^{i2}(1-q^i)^2 \Delta V^{i2}}$, and solve (2.5) for $\hat{\alpha}^i$. Then, fixing conjecture about population shares $\hat{q}^i$, for each $i \in I$ there are two conjectures about socialization efficiency, $\hat{\alpha}^i_h$ and $\hat{\alpha}^i_l$, that satisfy (2.5). They are given by the following[18]

$$\hat{\alpha}^i_h = \frac{\tau^i_{bv} + \xi^i}{2(1-\hat{q}^i)\Delta V^i}, \; \hat{\alpha}^i_l = \frac{\tau^i_{bv} - \xi^i}{2(1-\hat{q}^i)\Delta V^i} \tag{2.7}$$

For each $i \in I$, we can then compute the optimal socialization efforts. Since there are two possible equilibrium conjectures, there are also two possible equilibrium efforts. Define $\tau^i_h := \frac{\tau^i_{bv} + \xi^i}{2}$, and $\tau^i_l := \frac{\tau^i_{bv} - \xi^i}{2}$.

Next proposition characterizes the set of *selfconfirming* equilibria with respect different values of the conjectures $(\hat{q}^i)_{i \in I}$. Define by $E_i \in \mathcal{T} \times S$ the set of selfconfirming equilibrium choices and conjectures for each $i \in I$, with generic element $(\tau^i, \hat{\alpha}^i, \hat{q}^i)$. Then, the set of selfconfirming equilibria is given by $E := \times_{i \in I} E_i$.

Proposition 2 *For each $i \in I$:*

i) If $\hat{q}^i > \bar{q}^i$, it does not exist any $\hat{\alpha}^i \in [0,1]$ satisfying the confirmed conjectures property, and $E_i = \emptyset$;

ii) If $q^i < \hat{q}^i < \bar{q}^i$, $E_i = \{\{(\tau^i_h, \hat{\alpha}^i_h, \hat{q}^i)\}, \{(\tau^i_l, \hat{\alpha}^i_l, \hat{q}^i)\}\}$, with $\tau^i_l < \tau^i_h < \tau^i_{bv}$ and $\hat{\alpha}^i_l < \hat{\alpha}^i_h < \alpha^i$;

iii) If $\hat{q}^i = q^i$, $E_i = \{\{(\tau^i_{bv}, \alpha^i, q^i)\}\}$;

iv) If $\hat{q}^i < q^i < \bar{q}^i$, $E_i = \{\{(\tau^i_h, \hat{\alpha}^i_h, \hat{q}^i)\}\}$, with $\tau^i_h > \tau^i_{bv}$ and $\hat{\alpha}^i_h > \alpha^i$.

Proof. in the Appendix. □.

This proposition analyzes how efforts differ with respect to the Bisin Verdier framework, and how conjecture relates to the true parameter values. As a direct consequence of the multiplicity of confirmed conjectures discussed above, we observe in some cases a multiplicity of equilibria. This is discussed in Figure 2.2 that summarizes the results of Proposition 2, and in which the diagonal represents the equilibrium with correct conjectures (Bisin and Verdier, 2001). Start focusing on conjectures $\hat{q}^i$. If a group underestimates its own presence in the society $\hat{q}^i < q^i$

[18] There are some issues that we need to address. First of all some of these conjectures, while they satisfy (2.5), they may be unfeasible because negative. Second, even if feasible, each of the conjectures $(\hat{\alpha}_h, \hat{q}^i)$ and $(\hat{\alpha}_l, \hat{q}^i)$ induces a subjectively optimal vertical socialization effort that needs to be feasible. In the next proposition we address both these issues. Notice also that in some cases we can find that $\hat{\alpha}^i > 1$, that means that parents think that there are no frictions in the vertical transmission, but that, on the contrary, this transmission is extremely efficient. Most importantly, for every $\hat{q}^i < \bar{q}^i$, we have that $d^i \in [0,1]$.

the socialization choice of its members is unique and, due to a cultural substitution reasoning, always higher than in the baseline model Bisin and Verdier (2001) with complete information; on the other hand, if a group overestimate the presence of its own trait in the society, $\hat{q}^i > q^i$, there is room to two different conjectures and two different *selfconfirming* socialization choices, both below to the Bisin-Verdier benchmark, but with different magnitudes. This multiplicity derives from the multiplicity of confirmed conjectures discussed above. Notice also that in all cases, $\hat{q}^i < q^i$ if and only if $\hat{\alpha}^i > \alpha^i$. Then in equilibrium parents can overestimate own group in the society only underestimating their efficacy in the vertical socialization, and viceversa. This immediately comes from the functional form of the feedback function. This result can, at least in part, explain why the "Parental Self-Efficacy" are lower among immigrants. Agents belonging to minorities, in fact, has less room to underestimate their own presence in the society.
Consider now the case of $\hat{q}^i < q^i$ where the equilibrium is unique. To understand why, in this case, only one equilibrium is possible, notice that the conjecture $\hat{\alpha}^i_l$ would suggest to the parent to choose a negative socialization effort ($\tau^i < 0$) that is not possible. Therefore the parent would choose the corner solution ($\tau^i = 0$). However, given the shape of the feedback function, a null effort would be compatible just with a correct conjecture about q^i (since in this case $p^{ii} = q^i$), and this would contrast with the assumption that $\hat{q}^i < q^i$. Then, for the same reason, when $\hat{q}^i = q^i$ the Bisin Verdier result holds.

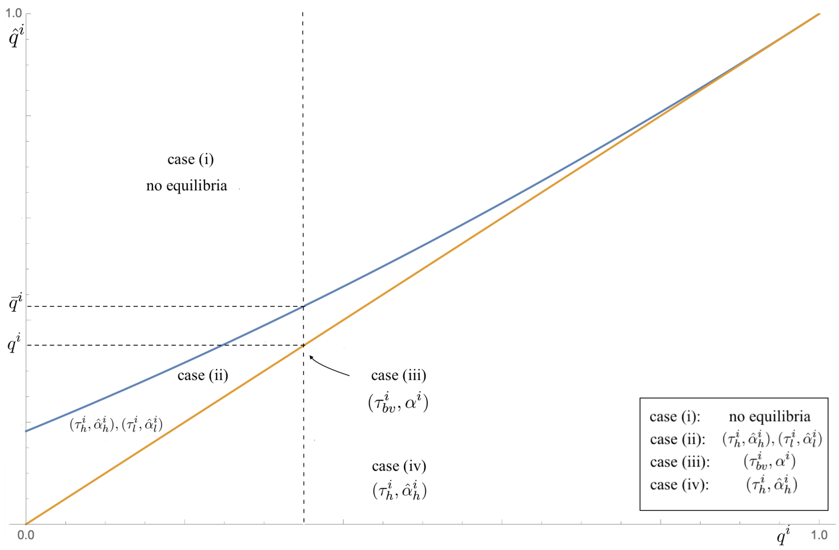


Figure 2.2: $\alpha = 0.9$, $\Delta V^i = 0.9$

2.3.1 Cultural Complementarity and Substitution

One of the main results of the standard models of cultural transmission with complete information is the *cultural substitution* property. This property requires that optimal socialization efforts are decreasing in own population shares. There is still debate about how to get cases in which the opposite property, *cultural complementarity*, holds. The two properties, then, have different consequences in terms of the derived population dynamics in particular for what regards the stability of polymorphic equilibria. In the case of complete information, cultural complementarity and substitution are uniquely and unambiguously defined since parents react to actual population shares, so that there is no difference between what agents think to be the social composition of the society affecting the oblique socialization process, and what this composition actually is. In the case of incomplete information, agents just have conjectures about population shares, so that efforts decision directly depends on the conjecture $\hat{q}^i$. On the other hand, q^i indirectly affects the socialization decision of agent i, since, given a certain conjecture $\hat{q}^i$, the confirmed conjecture $\hat{\alpha}^i$ depends on the true q^i, so that equilibrium τ^i depends on q^i itself in a non trivial way. Therefore it is worth defining the concept of cultural substitution or complementarity properties with respect to both the true parameter q^i and the conjecture $\hat{q}^i$.

Definition For each $i \in I$ and for each $\tau^i \in \{\tau_l^i, \tau_h^i\}$,

- τ^i displays ***actual* cultural substitution (complementarity)** if it is decreasing (increasing) in q^i.
- τ^i displays ***conjectured* cultural substitution (complementarity)** if it is decreasing (increasing) in $\hat{q}^i$.

Next proposition describes how the equilibrium socialization choice relates to actual and conjectured cultural complementarity and substitution.

Proposition 3 *For each $i \in I$, τ_h^i displays actual cultural complementarity and conjectured cultural substitution. τ_l^i displays actual cultural substitution and conjectured cultural complementarity.*

Proof. in the Appendix □.

It is possible to see results of Proposition 3 in Figure 3.2. In Figure 3.2a we can see how socialization choices relate to conjecture $\hat{q}^i$. If $\hat{q}^i < q^i$ the unique equilibrium socialization choice is negatively related to the conjecture about the share of traits in the society, so that there is *conjectured cultural substitution*. This result is not surprising since cultural substitution with respect to what people consider to be own population share in taking socialization choices is what we find in all model of cultural transmission. However, if $\hat{q}^i > q^i$ we have already seen that there are two possible equilibria and we observe both *cultural substitution* and *complementarity* associated respectively to the high and low socialization choice. In fact, if parents have

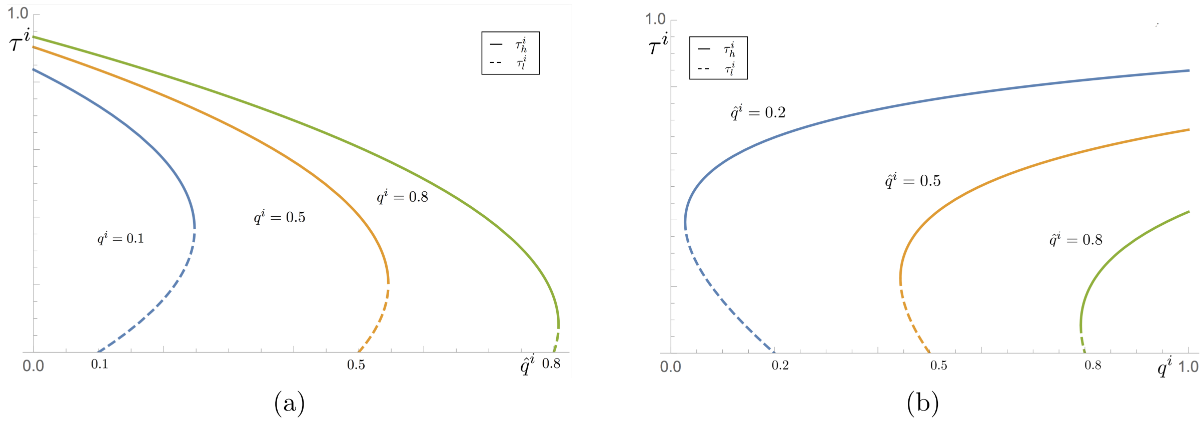


Figure 2.3: Socialization choice with respect: (a) $\hat{q}^i$ with $q^i = 0.1, 0.5, 0.9$, and (b) q^i with $\hat{q}^i = 0.2, 0.5, 0.8$, $\alpha = 0.9$. $\alpha = \Delta V^i = 0.9$

the the low efficiency conjecture $\hat{\alpha}^i$ then, when $\hat{q}^i > q^i$, they get stuck at the low socialization equilibrium and an increase in their conjecture leads to an increase in the socialization choice so that cultural complementarity is shown. The opposite occurs with the higher socialization effort.

Consider now the case for actual cultural complementarity and substitution. This is interesting since conjectures cannot be observed, so that the relationship between effort and actual population shares is what can be somehow observed. In this respect, Figure 3.2b shows that when $\hat{q}^i < q^i$ the unique equilibrium τ^i exhibits *cultural complementarity* property with respect to the true q^i. This result should be interpreted keeping in mind that, fixing the conjecture $\hat{q}^i$, the effect of q^i on the socialization choice is indirect and passes through the conjecture $\hat{\alpha}_h^i$. Using the learning interpretation of the equilibrium, if the presence of trait i in the society (q^i) increases, then parents can misinterpret the feedback they received and attribute the increase of p^{ii} to an increase on $\hat{\alpha}_h^i$, keeping fixed their conjecture $\hat{q}^i$, and thus they exert higher socialization effort ($\frac{\partial \tau^i}{\partial \hat{\alpha}_h^i} > 0$).

This makes particular sense because conjectures about population shares are more difficult to test and, thus, to revise for the individual in the short-run. Conjectures about population shares can be induced by social media, political or cultural leaders and news so that parents, by rational inattention, take them almost as given and produce compatible conjectures about what is the efficiency of the socialization process.

As briefly discussed in Section 3, if parents have correct conjectures about self-efficacy, $\hat{\alpha}^i = \alpha^i$, as in Bisin and Verdier (2001), then equilibria with wrong conjecture can exist if only if there is noisy signals about transition probabilities. In this case, however, parents have always cultural substitutions and not cultural complementarity.

2.3.2 Transition Probabilities

We have seen how incomplete information about the true parameters may produce multiple *selfconfirming* equilibria, the low and the high socialization equilibria. In this section we are interested in studying the transition probability p^{ii} stemming from *selfconfirming* equilibria and compare it with the benchmark Bisin-Verdier transition probability.

From the evaluation of the consequence function (2.1) at the two possible socialization efforts derived from confirmed conjectures, and substitution (2.7) in (2.1) we get

$$p_h^{ii*} = q^i + \alpha^i(1-q^i)\frac{\tau_{bv}^i + \xi^i}{2}, \quad p_l^{ii*} = q^i + \alpha^i(1-q^i)\frac{\tau_{bv}^i - \xi^i}{2} \tag{2.8}$$

Let p^{ii*} be the realized transition probability, then

Proposition 4

- *If* $\hat{q}^i < q^i$, $p^{ii*} = p_h^{ii}$ *with* $p^{ii} > p_{bv}^{ii}$
- *If* $\hat{q}^i > q^i$, $p^{ii*} \in \{p_h^{ii}, p_l^{ii}\}$ *with* $p_l^{ii} < p_h^{ii} < p_{bv}^{ii}$

Proof. in the Appendix □.

The presence of multiple equilibria when $\hat{q}^i > q^i$ induces a bifurcation in the population dynamics. We have, however, a reasonable way to select one of the two transition probabilities at each period. Recall that the presence of two transition probabilities is due to the presence of the two equilibrium efforts that, in turn, depends on the fact that for each conjecture $\hat{q}^i$ there are two feasible $\hat{\alpha}_h^i$ and $\hat{\alpha}_l^i$. Notice however that parents are not indifferent between these two conjectures. While they cannot choose between them in terms of which is the most likely, since the two conjectures are both confirmed by the feedback, they provide different levels of subjective expected utility, since they induce different effort levels and different transition probabilities. In particular the following result holds

Proposition 5 *For each* $i \in I$, *and for every triple* $(q^i, \alpha^i, \hat{q}^i)$ *such that the set of selfconfirming equilibria* $E_i = \{\{(\tau_h^i, \hat{\alpha}_h^i, \hat{q}^i)\}, \{(\tau_l^i, \hat{\alpha}_l^i, \hat{q}^i)\}\}$, $\mathbb{E}^i_{p_h^{ii}}[u^i(q^i)] > \mathbb{E}^i_{p_l^{ii}}[u^i(q^i)]$.

Proof. in the Appendix □.

Since parents know their subjective expected utility functional form, and since they want to maximize it, parents may be conscious that choosing $\hat{\alpha}_h^i$ as opposed to $\hat{\alpha}_l^i$, while being equally likely in terms of conjecture confirmation, provides a higher level of expected utility. Then, we assume that they choose the conjecture that maximizes their expected utility. In this way, when multiple equilibria are possibile, they always choose the one inducing higher effort and higher transition probability.

Assumption 1 *For each $i \in I$, if $q^i < \hat{q}^i$, $E_i = \{\{\tau_h^i, \hat{\alpha}_h^i, \hat{q}^i\}\}$.*

Notice that this equilibrium selection is corroborated by social psychological literature. Indeed, higher Parental Self-Efficiency is empirically associated with lower parental depression, anxiety, and stress, therefore higher parent's utility, and has positive effect on the development of the offspring (Jackson and Huang, 2000; Kuhn and Carter, 2006). Moreover, it has some properties that makes it particularly reasonable. In particular, it induces continuity in the effort with respect to conjectured population shares. Indeed, fixing a $\epsilon \in [0, 1]$ arbitrarily small, and consider $\hat{q}'^i := q^i - \epsilon$, then $\tau_h^i(\hat{q}'^i)$ is chosen. It is then likely that if parents experience a little increase in their conjectures, $\hat{q}''^i := q^i + \epsilon$, the effort choice is not very far from the previous one. Then $\tau_h^i(\hat{q}''^i)$ is more likely to be chosen than $\tau_l^i(\hat{q}''^i)$. Notice also that this equilibrium selection process is such conjectured cultural substitution is always shown. Again, this is a standard property of efforts with respect to what parents react to during the socialization process. This equilibrium selection also induces continuity of socialization effort on actual population shares, for the same reasoning explained above but applied to actual rather than conjectured population shared. At last, by Proposition **??** we aways observe actual cultural complementarity. This is a main difference with respect to literature about cultural transmission with complete information where parents always choose their socialization under cultural substitution. With incomplete information we show that parents think to choose the socialization effort under cultural substitutions, instead, since they do not know true population shares, the cultural substitutions property refers only to their conjecture. On the other hand, they actually chose under cultural complementarity, through the mechanism of confirmed conjectures.

2.4 Welfare with Wrong Conjectures

Before moving to the analysis of the population dynamics, we study how much the wrong but confirmed conjectures may cost in terms of utility loss. Notice that we refer only to a parent-specific type of welfare, in fact, parents are paternalistic and do not actually really care about the true welfare of their children. In details, whenever parents have correct conjectures they produce an effort that is optimal to the environment. The derived utility is our benchmark. However, when parents misperceived the population shares the equilibrium effort is ex-ante subjectively optimal but ex-post suboptimal. For each $i \in I$, define $\delta_{\hat{q}^i} := q^i - \hat{q}^i$. Notice however

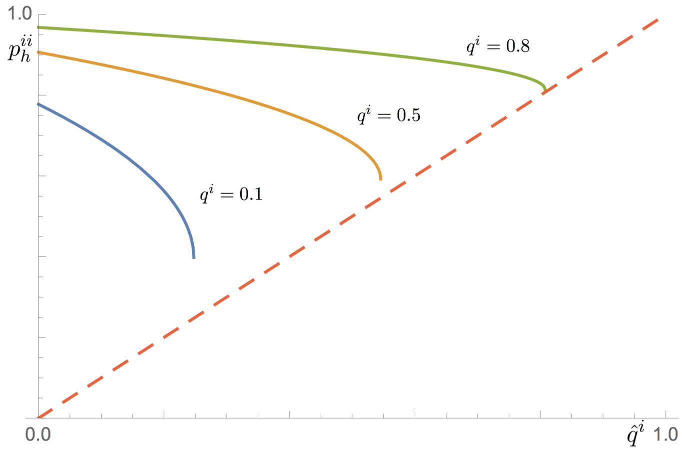


Figure 2.4: Transition probabilities with respect $\hat{q}^i$ with $q^i = 0.1, 0.5, 0.8$

that parents do not know the correct parameter q^i, otherwise they would have used it in the decision process. Then they cannot compute how much they loose in terms of expected utility and, for the same reason, how much they loose in terms of realized utility. A policymaker, however, may know the correct parameters, can compute both, and compute the loss in terms of ex post utility when agents use wrong conjectures.[19] Define $U^i_{bv} := U^i_{(\alpha^i,q^i)}(\alpha^i, q^i, \tau^i_{bv})$ as the average utility realized by i agents, at parameters (α^i, q^i) when they have correct conjectures and perform the Bisin-Verdier effort. In this case transition probabilities are computed at the correct q^i with the Bisin-Verdier efforts. Define $U^i_{\hat{q}^i} := U^i_{(\alpha^i,q^i)}(\hat{\alpha}^i_h, \hat{q}^i, \tau^i_h)$ as the average utility realized by i agents, at parameters (α^i, q^i) when they have conjectures $(\hat{\alpha}^i_h, \hat{q}^i)$ and exert effort τ^i_h. In this case transition probabilities are computed at the correct q^i and at the exerted effort τ^i_h. Let $\Delta U^i_{\hat{q}^i} := U^i_{bv} - U^i_{\hat{q}^i}$ be the average loss i agents experience having wrong conjectures. Then

Proposition 6 *For each $i \in I$, $\Delta U^i_{\hat{q}^i} = \frac{1}{8}(\tau^i_{bv} - \xi^i)^2 > 0$, and $\frac{\partial \Delta U^i_{\hat{q}^i}}{\partial \delta_{\hat{q}^i}} = \frac{\Delta V^i(\tau^i_{bv} - \xi^i)}{2\xi^i}$*

Proof. in the Appendix □

To inspect the properties of the loss it is worth looking at Figure 2.5 to see how it changes with respect to conjectures. The loss is a convex function of the distance from the correct value. In details, in the correct value it reaches, by definition, a null value. Interestingly, it is steeper for positive biases than for negative ones. This is due to the fact that a negative bias induces a higher effort than a positive bias, and the derived transmission probability, higher than the case of a positive bias, partially recover the loss. Another source of asymmetry is

[19] Notice that, while for an individual parent the realized utility depends on the realization of the socialization process, when talking about representative agents of a group, as we do in this paper, the utility function is exactly the average utility parents of type i gets, and the policy maker just looks at this.

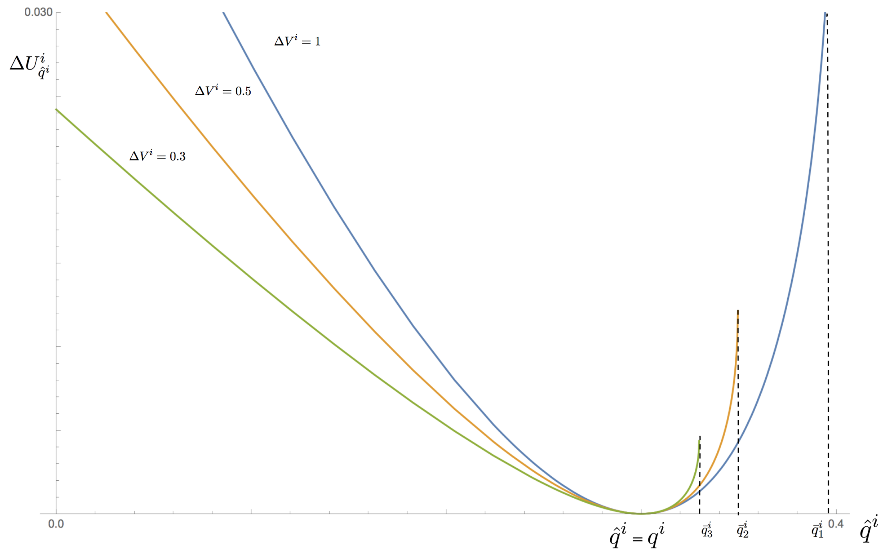


Figure 2.5: Loss

given by the different levels of ΔV^i. In details, given a downward bias, more intolerant agents experience a higher loss, while this relations is reverted for positive biases. Notice that, since the welfare is lower for each group, also the total welfare is lower. We can compute the welfare loss of the whole society given $\hat{q}^i$ and $\hat{q}^j$. This is given by $\frac{1}{8}[q^i(\tau^i_{bv} - \xi^i)^2 + (1 - q^i)(\tau^j_{bv} - \xi^j)^2]$.

These results have two main consequence. First of all, given a fixed misperception of own group population shares, agents have a lower loss by under perceiving that over perceiving the share. Second, whenever agents of a group under represent their own presence in the society, the more intolerant they are the more they loose in welfare, since the loss is increasing in ΔV^i.

It is important to underline how, in this framework, introducing the possibility for agents to pay a cost to receive better information does not affect the results. This is due to the fact that parents have deterministic conjectures and they are not aware that their conjecture can be wrong.[20] Thus, their marginal expected benefit to have new information is zero, even if their actual ex-post marginal benefit would be positive. To anticipate the discussion of the next section, if the signals are provided by cultural leaders, then agents may believe to leaders or have private different conjectures and believe in that. In both cases, they do not have incentives to pay to acquire new pieces of information.

[20]With probabilistic conjectures this results may be different and depend on parents' higher order beliefs.

2.5 Long-run Dynamics

In this section we study the consequences of incomplete information for the analysis of the long run dynamics of population traits. We find that incomplete information may totally revert standard prediction of cultural transmission literature. Incomplete information about population shares is a particularly relevant issue since conjectures about population shares can be shaped by social media, news or fake news, cultural leaders that may use this dimension when it comes to be a particularly salient issue in the public debate. We then introduce the presence of a cultural leader that induces one (possibly biased) conjectures in own group. On the other hand, conjecture about parental efficacy is strictly related to the parent-offspring relationship and then we assume that each parent has some (possibly wrong) conjecture that cannot be shaped by the policy maker. We discuss how the bias in the population shares conjectures induced by the cultural leaders may drastically change the long run dynamics of the population. We introduce time indexes for all the quantities. In details, for each time t, the equilibrium transition probabilities are given by p_t^{ii*} and p_t^{jj*}. Notice that equilibrium transition probabilities are defined given conjectures $\hat{q}^i$ and $\hat{q}^j$ and the derived α_h^i and α_h^j. In Appendix 2.6 we show what happens if conjectures about population shares do not evolve along time and prove that this may not be sustainable in the long run with these conjectures being confirmed.

We first set the dynamic system. Recalling that, for each $i \in I$, $d^{i*} = \alpha^i \tau^{i*}$, for every given time t, the population share of type $i \in I$ in the subsequent period is described by the following law

$$q_{t+1}^i = p_t^{ii*} q_t^i + p_t^{ji*}(1 - q_t^i) \tag{2.9}$$

$$= q_t^i[1 + (d_t^{i*} - d_t^{j*})(1 - q_t^i)] \tag{2.10}$$

Using a continuous time approximation we get

$$\dot{q}_t^i = q_t^i(1 - q_t^i)(d_t^{i*} - d_t^{j*}) \tag{2.11}$$

At this stage, we introduce a cultural leader for each community, l^i, l^j.[21] Leaders have the power of inducing a bias p (positive) or n (negative) in the perception of their community members, but cannot control the intensity of biases, which we consider exogenous and constant along all the dynamics. For each $i \in I$, intensity of the bias if given by $\beta^i \in (0, 1)$.[22] Thus leader l^i chooses an action $a_i \in \{p, n\}$.

[21] Among some previous work about cultural leaders Nteta and Wallsten (2012), Acemoglu and Jackson (2014),Verdier and Zenou (2015), Prummer and Siedlarek (2017), Verdier and Zenou (2018)

[22] Endogenizing the intensity of the bias β^i does not bring any meaningful insight.

Define an indicator function $\mathbb{I}_{a^i}$ which takes value 0 if the leader i chooses $a^i = n$ and 1 if she chooses $a^i = p$. Then we define biased conjectures for community i as

$$\hat{q}^i_{\beta,t} = \beta^i \mathbb{I}_{a^i} + (1-\beta^i) q^i_t \tag{2.12}$$

Thus, if a leader chooses a positive bias the derived conjecture about population share will be $\hat{q}^i_t = \beta + (1-\beta)q^i_t > q^i_t$; if the leader choses a negative bias, then $\hat{q}^i_t = (1-\beta)q^i_t < q^i_t$. Both leaders, at each time t, are not forward looking[23] and are interested only in maximizing the presence of their trait in the society one period ahead. Therefore

$$\max_{a^i} \; u^{l^i}_t(a_i, a_j) = q^i_{t+1}(a_i, a_j) \tag{2.13}$$

Proposition 7 *A leader who faces the maximization problem (2.13) chooses always $a_i = n$. Then $\hat{q}^i_{\beta,t} = (1-\beta^i)q^i_t$.*

Proof. in the Appendix □

Following Proposition **??**, cultural leaders, in order to maximize the share of the trait of her own community in the next period, always choose to instill a negative bias in agents belonging to their cultural group. This is a simple results that derives from the fact that, independently on other leader's choice, it is always better to induce a negative bias in own population conjectures, since this would induce a higher socialization effort and a higher share of own type in next generation, so that for each leader $i \in I$, choosing a negative bias is a dominant strategy.

We have assumed that cultural leaders are myopic in their decisions to choose their group's bias, thus we implicitly assume that leaders are individuals with a finite life. We can also think about leaders as centralized cultural institutions with a more forward looking perspective, in that case, the game between cultural leaders is a differential dynamic game. However, results do not change. The space of action is dichotomous, positive or negative bias. Moreover, as shown in Section 3.1, with negative biases there is always cultural complementarity in socialization effort, thus a negative bias leads to a higher parental effort. For these reasons, for cultural leaders maximizing the presence of their trait in the society is equivalent to maximize to long-run one.

[23]Since we are in an intergenerational setting, leaders cannot live forever.

We can now consider the long-run dynamics of (2.22) with biased conjectures.

$$\dot{q}_t^i = \frac{1}{2} q_t^i (1 - q_t^i) \phi(q_t^i, \hat{q}_{\beta,t}^i, \hat{q}_{\beta,t}^j) \tag{2.14}$$

Before providing a characterization of steady states, define the following thresholds:

$$\bar{\beta}^i := \frac{\alpha^{j4}}{\alpha^{i2}} \frac{\Delta V^{j2}}{\Delta V^i}, \ \bar{\beta}^j := \frac{\alpha^{i4}}{\alpha^{j2}} \frac{\Delta V^{i2}}{\Delta V^j} \tag{2.15}$$

Proposition 8 *Given $a_i = n$, the dynamics (2.14) is well defined for all $q^i \in [0,1]$ for all q^i. $\{0,1\} \subseteq Q_{ss}^i$ and there is at most one polymorphic steady state $q^* \in (0,1)$. Moreover, for each $i \in I$,*

- *$q^i = 1$ ($q^i = 0$) is stable if and only if $\beta^i > \bar{\beta}^i$ ($\beta^j > \bar{\beta}^j$)*
- *q^* is stable if and only if $\beta^i < \bar{\beta}^i$ and $\beta^j < \bar{\beta}^j$*

Proof. in the Appendix □

Proposition 8 states that, depending on primitive parameters of the model (α and ΔV) and bias β, the cultural dynamics can show both stable and unstable polymorphic equilibria, and also no polymorphic equilibria at all.
Consider Figure 2.6 that describes all the possibile cases that can happen in the dynamics. If the negative biases are both large (part I of the graph), then an unstable polymorphic equilibrium is shown. This derives from the fact that for large negative biases, each group thinks to be extremely smaller than what it is in reality. Then, as we have already seen in previous section, while we have conjectured cultural substitution, our model predicts actual cultural complementarity, and then unstable polymorphic equilibrium is observed. The opposite occurs for small negative biases (part III of the graph). In this case agents have almost correct conjectures, and then Bisin Verdier results are good proxies of the dynamics. If, on the contrary, the biases are unbalanced, then one group takes over the other and invades the society.

Notice that the presence of groups' heterogeneity in, at least one, between the biases or in α and ΔV produces the possibility of globally stable homomorphic equilibria. Consider, for example, the case of Figure 2.6 when $\beta^i = \beta^j$. The dynamics can lay in parts I, III and IV of the graph. This is due to the difference between the intolerance of agents belonging to the two groups. On the contrary, if there is complete homogeneity in the α, ΔV and β then the dynamics have always, stable or unstable, polymorphic equilibria (I and III in Figure 6).

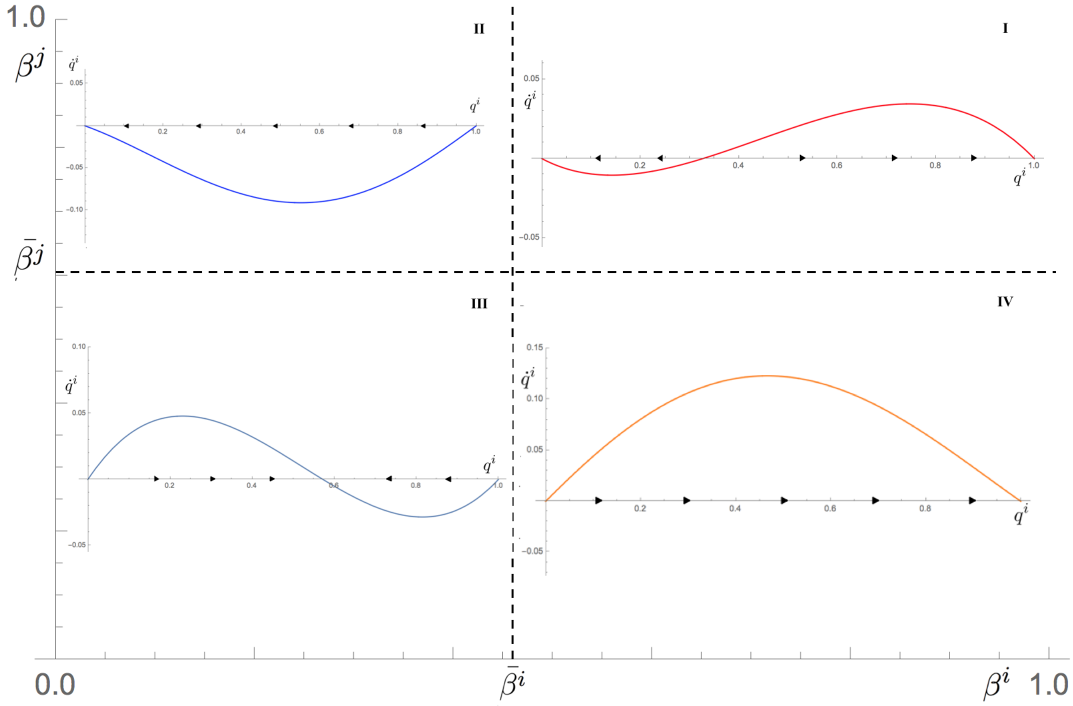


Figure 2.6: Cultural dynamics $\dot{q}^i$ for different β^i, β^j with $\alpha^i = \alpha^j = 1$, $\Delta V^i = 0.55$, $\Delta V^j = 0.5$.

This result has important implications both from a theoretical point of view and from an applied prospective.

Let us focus first on the difference between the population dynamic with incomplete information as compared to the case in which agents have complete information. In the standard case, as (Bisin and Verdier, 2001; Cheung and Wu, 2018), cultural substitution always leads to cultural heterogeneity while cultural complementarity leads to cultural homogeneity. In an incomplete information setting, however, in spite of having actual cultural complementarity in socialization choice, under certain condition, there exists a stable equilibrium with cultural heterogeneity. Interestingly, the primitives of the model $(\Delta V^i, \alpha^i)_{i \in I}$ determine the thresholds $\bar{\beta}^i$ and $\bar{\beta}^j$ and thus the different long run that are possible given reciprocal intolerance and actual parental efficacy. Then, the actual biases β^i and β^j determine which of the four possible outcomes is selected.

We know move to discuss applied implication of results. Let us consider a society where one of the two leaders is identitarian and maximize (2.13) and a non-identitarian, or group utilitarian, leader that maximize welfare. The identitarian leader decides to induce a negative bias, while the non-identitarian truthfully report the true population shares. In such a situation, looking at Figure 2.6 we move only in one axis. Assuming i is the non identitarian leader, then $\beta^i = 0$, thus the only two possible social outcomes are I and III, namely depending on β^j a stable polymorphic equilibrium or globally stable homomorphic equilibrium where dominate j can exist. In particular, we can see that the more the agents belonging to the group with identitarian leader intolerant (ΔV^j) or effective as parents (α^j) are, the more is the space for domination of that group in the long run. The same reasoning applies to the case with two identitarian leaders, where if the $\beta^i = \beta^j$, the intolerance and parental efficacy of the two groups determine population shares in the equilibrium with cultural heterogeneity. Moreover, starting from III, where $\beta^i < \bar{\beta}^i$ and $\beta^j < \bar{\beta}^j$ and keeping β^j fixed, the more β^i grows and approaches $\bar{\beta}^i$ the more the polymorphic equilibrium moves to the right increasing the share of agents belonging to i in the society.

2.6 Conclusion

This paper generalizes the classical cultural transmission model proposed by Bisin and Verdier (2001) to an incomplete information setting. We show that if parents are not fully aware of own group size and about the efficiency of the cultural transmission technology, they may end up to sustain wrong conjectures about both quantities. While in the standard setting there is no difference between conjectures and true parameters, with incomplete information this difference is crucial. Thus, we propose the definition of "conjectured cultural substitution (complementarity") as opposed to the "actual cultural substitution (complementarity"). The main finding is that, with incomplete information, it is possible to obtain "actual cultural complementarity" in the standard cultural transmission mechanism. In the long-run we are able to reproduce all the possible social outcome as depending on the magnitude of bias on the population share. The most interesting result is that, in spite of having actual cultural complementarity in socialization choice, can exist, in the long run, a stable equilibrium with cultural heterogeneity. This result is particularly relevant because in previous cultural transmission papers, cultural complementarity leads always to cultural homogeneity.

Appendix A

Proofs of Propositions

Proof of Proposition 1

Solving the problem in (2.3) we can see that *subjective rationality* is satisfied by (2.4)

$$\tau^i = \hat{\alpha}^i(1-\hat{q}^i)\Delta V^i$$

Where $\Delta V^i = V^{ii} - V^{ij}$ represent the *cultural intolerance* of a parent of trait i.

To have *selfconfirming* equilibria parent's conjectures, $\hat{\alpha}^i$ and $\hat{q}^i$, have to be *confirmed* by the experience (feedback). Assuming m^i, the message of feedback function of agent i, being equal to p^{ii}, namely $f = g$, then the conjectures are confirmed if and only if

$$\alpha\tau^i + (1-\alpha\tau^i)q^i = \hat{\alpha}\tau^i + (1-\hat{\alpha}\tau^i)\hat{q}^i$$

therefore the conjectures that support *selfconfirming* equilibria are described by (2.5)

$$(\hat{\alpha}^i, \hat{q}^i): \qquad \hat{\alpha}^i(1-\hat{q}^i)\Delta V^i(\alpha(1-q^i) - \hat{\alpha}(1-\hat{q}^i)) = \hat{q}^i - q^i$$

□

Proof of Proposition 2

From equation (2.5) we get the following condition on $\hat{\alpha}^i$ and $\hat{q}^i$

$$\hat{\alpha}^i\alpha^i(1-q^i)(1-\hat{q}^i)\Delta V^i - \hat{\alpha}^{i2}(1-\hat{q}^i)^2\Delta V^i + (q^i - \hat{q}^i) = 0$$

Solving for $\hat{\alpha}^i$ we get the conjectures on α^i that support *selfconfirming* equilibria

$$\hat{\alpha} = \frac{\alpha(1-q^i)\sqrt{\Delta V^i} \pm \sqrt{4(q^i - \hat{q}^i) + \alpha^2(1-q^i)^2\Delta V^i}}{2(1-\hat{q}^i)\sqrt{\Delta V^i}} \tag{2.16}$$

$\hat{\alpha}^i \; \exists \;\; iff \;\; \xi^i \; \geq 0$, namely $\hat{q}^i \leq \bar{q}^i$, with $\bar{q}^i = q^i + \frac{\alpha^{i2}(1-q^i)^2\Delta V^i}{4}$.

$$if \quad \hat{q}^i > \bar{q}^i = q^i + \frac{\alpha^{i2}(1-q^i)^2\Delta V^i}{4} \Rightarrow \nexists \; \hat{\alpha}^i \tag{2.17}$$

Substituting (2.16) in (2.4) we get

$$\tau^i = \frac{\alpha^i(1-q^i)\Delta V^i \pm \sqrt{4(q^i-\hat{q}^i)\Delta V^i + \alpha^{i2}(1-q^i)^2\Delta V^{i2}}}{2} \tag{2.18}$$

Let call $\xi^i = \sqrt{4(q^i-\hat{q}^i)\Delta V^i + \alpha^{i2}(1-q^i)^2\Delta V^{i2}}$ we can write the previous equation as

$$\tau^i = \frac{\tau^i_{bv} \ \pm \ \xi^i}{2} \tag{2.19}$$

Where $\xi^i \lesseqgtr \tau^i$ if $q^i \lesseqgtr \hat{q}^i$.

Since the existence of τ^i depends on the existence of $\hat{\alpha}^i$ we can say that

$$if \quad \hat{q}^i > \bar{q}^i = q^i + \frac{\alpha^{i2}(1-q^i)^2\Delta V^i}{4} \Rightarrow E^i = 0 \tag{2.20}$$

Moreover

$$\tau^i = \frac{\alpha^i(1-q^i)\Delta V^i \pm \sqrt{4(q^i-\hat{q}^i)\Delta V^i + \alpha^{i2}(1-q^i)^2\Delta V^{i2}}}{2} \in [0,1]$$

Defining

$$\tau^i_h = \frac{\tau^i_{bv} \ + \ \xi^i}{2}, \quad \tau^i_h = \frac{\tau^i_{bv} \ - \ \xi^i}{2} \tag{2.21}$$

It is trivial to see that τ^i_h is always positive.

We can further prove that $\tau^i_h = \frac{\tau^i_{bv} \ + \ \xi^i}{2} \leq 1$ always, in fact

$$\frac{\sqrt{4(q^i-\hat{q}^i)\Delta V^i + \alpha^{i2}(1-q^i)^2\Delta V^{i2}}}{2} \leq 1 - \frac{\tau^i_{bv}}{2}$$

$$\Rightarrow \ \sqrt{4(q^i-\hat{q}^i)\Delta V^i + \alpha^{i2}(1-q^i)^2\Delta V^{i2}} \leq 2 - \tau^i_{bv}$$

$$\Rightarrow \ 4(q^i-\hat{q}^i)\Delta V^i + \alpha^{i2}(1-q^i)^2\Delta V^{i2} \leq 4 + \tau^{i2} - 4\tau^i_{bv}$$

$$\Rightarrow \ 4(q^i-\hat{q}^i)\Delta V^i \leq 4 - 4\tau^i_{bv}$$

$$\Rightarrow \ (q^i-\hat{q}^i)\Delta V^i \leq 1 - \tau^i_{bv}$$

$$\Rightarrow \ \hat{q}^i - q^i \geq \underbrace{\frac{1}{2}\alpha(1-q^i)}_{\leq\frac{1}{2}} - \underbrace{\frac{1}{\Delta V^i}}_{\geq 1} \quad \text{Always satisfied!!}$$

$$\Rightarrow \ \tau_h^i \leq 1$$

With the same reasoning we verify under which condition $\tau_l^i \in [0,1]$.

Since $\tau_l^i \leq \tau_h^i \leq 1$ it is enough to verify conditions for $\tau_l^i > 0$. Given the definition τ_l^i is positive if and only if

$$\alpha(1-q^i)\sqrt{\Delta V^i} \geq \sqrt{4(q^i - \hat{q}^i) + \alpha^2(1-q^i)^2\Delta V^i}$$

$$\Rightarrow \ \alpha^2(1-q^i)^2\Delta V^i \geq 4(q^i - \hat{q}^i) + \alpha^2(1-q^i)^2\Delta V^i$$

$$\Rightarrow \ \hat{q}^i \geq q^i$$

We can now sum up the above results and obtain exactly the conditions of Proposition 2. All the conditions regarding the relation ship between α^i, $\hat{\alpha}_h^i$ and $\hat{\alpha}_l^i$ stem from (2.16). The first point of Proposition 2 is proven by (2.17) and (2.20). Then, we have shown that if $\hat{q}^i \leq \bar{q}^i$ then τ_h^i always exists and belongs to $[0,1]$ and that τ_l^i is positive if and only if $\hat{q}^i \geq q^i$. In the end, it is trivial to see by (2.18) that whenever $\hat{q}^i = q^i$ then $\tau^i = \tau_{bv}^i$ and that $\tau_{bv}^i \leq \tau_h^i$ always.

□

Proof of Proposition 3

In order to prove Proposition 3 we should consider the sign of first derivative of τ_h^i and τ_l^i with respect both $\hat{q}^i$ and q^i. Before to proceed we recall from Proposition 2 that if $\hat{q}^i < q^i$ then $\exists\ \tau_h^i$ but $\nexists\ \tau_l^i$, while if $\hat{q}^i > q^i$ then $\exists\ \tau_h^i, \tau_l^i$

- **Conjectured cultural substitution (complementarity)**

$$\frac{\partial \tau_h^i}{\partial \hat{q}^i} = -\frac{\Delta V^i}{\sqrt{4(q^i - \hat{q}^i) + \alpha^{i2}(1-q^i)^2\Delta V^i}} < 0$$

$\Rightarrow$ Cultural Substitution

$$\frac{\partial \tau_l^i}{\partial \hat{q}^i} = \frac{\Delta V^i}{\sqrt{4(q^i - \hat{q}^i) + \alpha^{i2}(1-q^i)^2\Delta V^i}} > 0$$

$\Rightarrow$ Cultural Complementarity

- **Actual cultural substitution (complementarity)**

$$\frac{\partial \tau_h^i}{\partial q^i} = \frac{1}{2}\Delta V^i \left(\frac{2 - \alpha^{i2}(1-q^i)\Delta V^i}{\sqrt{\Delta V^i(4(q^i - \hat{q}^i) + \alpha^{i2}(1-q^i)^2\Delta V^i)}} - \alpha \right) > 0$$

$$2 - \alpha^{i2}(1-q^i)\Delta V^i - \alpha\sqrt{\Delta V^i(4(q^i - \hat{q}^i) + \alpha^{i2}(1-q^i)^2\Delta V^i)} \geq 0$$

Since this function is always decreasing in both α^i and ΔV^i and increasing in $\hat{q}^i$ then it reach its minimum at $\alpha^i = \Delta V^i = 1$ and $\hat{q}^i = 0$, namely

$$2 - (1-q^i) - \sqrt{4q^i + (1-q^i)^2} \geq 0$$

Therefore there is cultural complementarity, if

$$2 \geq (1-q^i) + \sqrt{q^{i2} + 2q^i + 1+}$$

$$2 \geq (1-q^i) + (1+q^i) = 2$$

Therefore if $\alpha^i = \Delta V^i = 1$ and $\hat{q}^i = 0$ then $\frac{\partial \tau_h^i}{\partial q^i} = 0$ otherwise, $\frac{\partial \tau_h^i}{\partial q^i} > 0$

$\Rightarrow$ Cultural Complemetarity

$$\frac{\partial \tau_h^i}{\partial q^i} = \frac{1}{2}\Delta V^i \left(-\alpha - \frac{2 - \alpha^{i2}(1-q^i)\Delta V^i}{\sqrt{\Delta V^i(4(q^i - \hat{q}^i) + \alpha^{i2}(1-q^i)^2\Delta V^i)}} \right) < 0$$

$\Rightarrow$ Cultural Substitution

□

Proof of Proposition 4

$$p^{ii*} = d^i + (1-d^i)q^i$$

$$p^{ij*} = 1 - p^{ii}$$

$$p^{ii*} = \alpha^i\tau^i + (1-\alpha^i\tau^i)q^i$$

$$p^{ii*} = q^i + \tau^i\alpha^i(1-q^i)$$

$$p^{ii*} = q^i + \frac{\tau_{bv}^i \pm \xi^i}{2}\alpha^i(1-q^i)$$

by Proposition 2 we know that if $\hat{q}^i < q^i$ then $\tau_l^i = \frac{\tau_{bv}^i - \xi^i}{2}$ do not exists, thus p^{ii*} is unique and with $\hat{q}^i > q^i$ there are two p^{ii*}: (i) p_h^{ii*} associated to τ_h^i and (ii) p_l^{ii*} associated to τ_l^i.

Moreover, p^{ii*} is increasing in τ^i and we know that $\tau_h^i > \tau_{bv}^i > \tau_l^i$, therefore

- If $\hat{q}^i < q^i$ then $\tau^i = \tau_h^i > \tau_{bv}^i$, therefore $p^{ii*} = p_h^{ii*} > p_{bv}^{ii*}$ and
- If $\hat{q}^i > q^i$ then $\tau^i = \{\tau_l^i, \tau_h^i\} < \tau_{bv}^i$, therefore $p^{ii*} \in \{p_l^{ii*}, p_h^{ii*}\} < p_{bv}^{ii*}$

□

Proof of Proposition 5

Substituting τ_h^i and τ_l^i in the expected utility function we get

$$\mathbb{E}[u^i(\hat{q}^i, \tau_h^i)] = V^{ij} + \Delta V^i \left(\hat{q}^i + (1 - \hat{q}^i)\hat{\alpha}^i \frac{\tau_{bv}^i + \xi^i}{2} \right) - \frac{1}{2} \left(\frac{\tau_{bv}^i + \xi^i}{2} \right)^2$$

and

$$\mathbb{E}[u^i(\hat{q}^i, \tau_l^i)] = V^{ij} + \Delta V^i \left(\hat{q}^i + (1 - \hat{q}^i)\hat{\alpha}^i \frac{\tau_{bv}^i - \xi^i}{2} \right) - \frac{1}{2} \left(\frac{\tau_{bv}^i - \xi^i}{2} \right)^2$$

The difference is always greater than zero

$$\mathbb{E}[u^i(\hat{q}^i, \tau_h^i)] - \mathbb{E}[u^i(\hat{q}^i, \tau_l^i)] = \hat{\alpha}^i \Delta V^i \xi^i \left(1 - \hat{q} - \frac{1}{2}(1 - q) \right) \geq 0$$

if $\frac{1}{2} - \hat{q}^i + \frac{1}{2}q^i \geq 0 \;\Rightarrow\; \hat{q}^i < \frac{1}{2}(1 + q^i)$ that is always satisfied, since $\frac{1}{2}(1 + q^i) \geq \bar{q}^i$ and $\hat{q}^i \leq \bar{q}^i$

□

Proof of Proposition 6

The utility gained under complete information is

$$u^i(q^i, \tau_{bv}^i) = V^{ij} + \Delta V^i \left(q^i + (1 - q^i)\alpha^i \tau_{bv}^i \right) - \frac{1}{2} \left(\tau_{bv}^i \right)^2$$

Let us call

$$\Delta U_{\hat{q}^i}^i = u^i(q^i, \tau_{bv}^i) - u^i(q^i, \tau_h^i)$$

$$\Rightarrow \Delta U_{\hat{q}^i}^i = V^{ij} + \Delta V^i \left(q^i + (1 - q^i)\alpha^i \tau_{bv}^i \right) - \frac{1}{2} \left(\tau_{bv}^i \right)^2 - V^{ij} - \Delta V^i \left(q^i + (1 - q^i)\alpha^i \frac{\tau_{bv}^i + \xi^i}{2} \right) + \frac{1}{2} \left(\frac{\tau_{bv}^i + \xi^i}{2} \right)^2$$

$$\Rightarrow \Delta U^i_{\hat{q}^i} = \Delta V^i(1-q^i)\alpha^i\tau^i_{bv} - \frac{1}{2}\left(\tau^i_{bv}\right)^2 - \Delta V^i(1-q^i)\alpha^i\frac{\tau^i_{bv}+\xi^i}{2} + \frac{1}{2}\left(\frac{\tau^i_{bv}+\xi^i}{2}\right)^2$$

$$\Rightarrow \Delta U^i_{\hat{q}^i} = \tau^{2i}_{bv} - \frac{1}{2}\left(\tau^i_{bv}\right)^2 - \tau^i_{bv}\frac{\tau^i_{bv}+\xi^i}{2} + \frac{1}{2}\left(\frac{\tau^i_{bv}+\xi^i}{2}\right)^2$$

$$\Rightarrow \Delta U^i_{\hat{q}^i} = \frac{1}{2}\left(-\tau^i_{bv}\xi^i + \left(\frac{\tau^i_{bv}+\xi^i}{2}\right)^2\right)$$

$$\Rightarrow \Delta U^i_{\hat{q}^i} = \frac{1}{2}\left(-\tau^i_{bv}\xi^i + \frac{\tau^{i2}_{bv}+\xi^{i2}+2\tau^i_{bv}\xi^i}{4}\right)$$

$$\Rightarrow \Delta U^i_{\hat{q}^i} = \frac{1}{8}(\tau_{bv}-\xi^i)^2 > 0 \qquad \forall \alpha^i, \Delta V^i, q^i, \hat{q}^i$$

□

Proof of Proposition 7

$$q^i_{t+1} = p^{ii*}_t(a_i)q^i_t + (p^{jj*}_t(a_j))(1-q^i_t)$$

Since $p^{ii*}_t(a_i)$ is the only element of q^i_{t+1} that depends on a_i and $\frac{q^i_{t+1}}{p^{ii*}_t(a_i)} > 0$ then

$$\max_{a^i}\ q^i_{t+1}(a_i, a_j) = \max_{a^i}\ p^{ii}_t(a_i)$$

Given assumption 1, the transition probability of each period the only transition probability is

$$p^{ii}_t = p^{ii}_{h,t} = q^i_t + \frac{\tau^i_{bv,t} + \xi^i_t(\hat{q}^i_{\beta,t})}{2}\alpha^i_t(1-q^i_t)$$

Defining p^{ii}_- the transition probability associated with a negative biased $\hat{q}^i$ and p^{ii}_+ the transition probability associated with a positive biased $\hat{q}^i$ and since $\xi^i \lesseqgtr \tau^i_{bv}$ iff $q^i \lesseqgtr \hat{q}^i$ then we can state that $p^{ii}_- > p^{ii}_+$. Therefore $a_i = n$ (and $a_j = n$) is a dominant strategy for leader l^i (and l^j).

□

Proof of Proposition 8

It is possible to express the condition for existence of *selfconfirming* equilibria, $\hat{q}^i_t < \bar{q}^i_t = q^i_t + \frac{\tau^{i2}_{bv,t}}{4\Delta V^i}$ with respect β. Since the bias is always negative the condition becomes

$$(1-\beta^i)q^i_t < q^i_t + \frac{\tau^{i2}_{bv,t}}{4\Delta V^i}$$

$$\Rightarrow \beta^i > -\frac{\tau_{bv,t}^{i2}}{4\Delta V^i q_t^i}$$

Which is always satisfied, therefore $\dot{q}^i$ does exist for all q^i. $\square$

Notice that

$$\frac{\partial \dot{q}}{\partial q^i} = \frac{1}{2}(q^i(1-q^i)\frac{\partial \phi}{\partial q^i} + (1-2q^i)\phi)$$

therefore

$$\frac{\partial \dot{q}}{\partial q^i}|_{q^i=0} = \phi \;\text{ and }\; \frac{\partial \dot{q}}{\partial q^i}|_{q^i=1} = \frac{1}{2}\phi$$

Thus, to study the sign of $\frac{\partial \dot{q}}{\partial q^i}$ it is enough to look at ϕ:

$$\text{if }\; \phi|_{q^i=0} > 0 \quad \Rightarrow \quad \frac{\partial \dot{q}}{\partial q^i}|_{q^i=0} > 0$$

$$\text{if }\; \phi|_{q^i=1} > 0 \quad \Rightarrow \quad \frac{\partial \dot{q}}{\partial q^i}|_{q^i=1} > 0$$

.

$$\xi_i = \sqrt{4(q^i - (1-\beta^i)q^i)\Delta V^i + \alpha^{i2}(1-q^i)^2\Delta V^{i2}}$$
$$\xi_j = \sqrt{4(1-q^i-(1-\beta^j)(1-q^i))\Delta V^j + \alpha^{j2}q^{i2}\Delta V^{j2}}$$

$$\phi(q^i, \hat{q}_\beta^i, \hat{q}_\beta^j) = \alpha^{i2}(1-q^i)\Delta V^i + \alpha^i\xi^i(\hat{q}_\beta^i) - \alpha^{j2}q^i\Delta V^j - \alpha^j\xi^j(\hat{q}_\beta^j)$$

-
$$\phi(q^i, \hat{q}_\beta^i, \hat{q}_\beta^j)|_{q^i=0} = \alpha^{i2}\Delta V^i + \alpha^i\xi^i|_{q^i=0}(\hat{q}_\beta^i) - \alpha^j\xi^j|_{q^i=0}(\hat{q}_\beta^j)$$
$$\phi(q^i, \hat{q}_\beta^i, \hat{q}_\beta^j)|_{q^i=0} = 2\alpha^{i2}\Delta V^i - 2\alpha^j\sqrt{\beta^j\Delta V^j} = 0$$
$$\sqrt{\beta^j} = \frac{\alpha^{i2}}{\alpha^j}\frac{\Delta V^i}{\sqrt{\Delta V^j}}$$
$$\beta^j = \frac{\alpha^{i4}}{\alpha^{j2}}\frac{\Delta V^{i2}}{\Delta V^j}$$
$$\beta^j \leq \frac{\alpha^{i4}}{\alpha^{j2}}\frac{\Delta V^{i2}}{\Delta V^j} \quad \Rightarrow \quad \phi(q^i, \hat{q}_\beta^i, \hat{q}_\beta^j)|_{q^i=0} \geq 0$$

-
$$\phi(q^i, \hat{q}_\beta^i, \hat{q}_\beta^j)|_{q^i=1} = \alpha^i\sqrt{4\beta^i\Delta V^i} - \alpha^{j2}\Delta V^j - \alpha^{j2}\Delta V^j$$

$$\sqrt{\beta^i} = \frac{\alpha^{j2}}{\alpha^i}\frac{\Delta V^j}{\sqrt{\Delta V^i}}$$

$$\beta^i = \frac{\alpha^{j4}}{\alpha^{i2}}\frac{\Delta V^{j2}}{\Delta V^i}$$

$$\beta^i \leq \frac{\alpha^{j4}}{\alpha^{i2}}\frac{\Delta V^{j2}}{\Delta V^i} \quad \Rightarrow \quad \phi(q^i, \hat{q}^i_\beta, \hat{q}^j_\beta)|_{q^i=1} \geq 0$$

□

Appendix B

The case for time invariant conjectures

Consider the case in which, for each $i \in I$, conjectures $\hat{q}^i$ are time invariant. After a little algebra, substituting equilibrium quantities into (2.11), we get

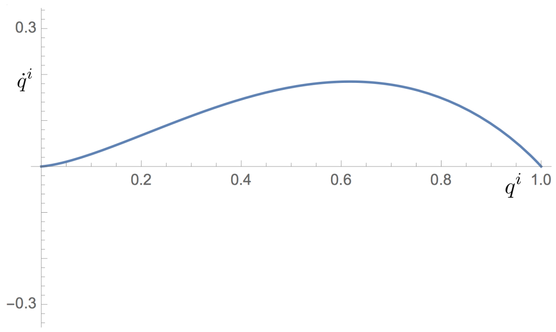
$$\dot{q}^i_t = \frac{1}{2}q^i_t(1-q^i_t)\phi(q^i_t, \hat{q}^i, \hat{q}^j) \tag{2.22}$$

where $\phi(q^i_t, \hat{q}^i, \hat{q}^j) := (\alpha^i)^2(1-q^i_t)\Delta V^i + \alpha^i\xi^i_t - (\alpha^j)^2 q^i_t \Delta V^j - \alpha^j_t\xi^j_t$. Recall that, even if we consider $\hat{q}^i_t$ as fixed, however the selfconfirming equilibrium efforts and conjectures are defined as long as, for each $i \in I$, and for each time t, $\hat{q}^i < \bar{q}^i_t$. Indeed, the threshold $\bar{q}^i_t$ depends on q_t and thus varies with time. It is then possibile that, for a given conjecture $\hat{q}^i_t$ the population dynamics evolve such that at some point the existence condition is not satisfied anymore. Recall that $\bar{q}^i_t := q^i_t + \frac{{\alpha^i}^2\Delta V^i}{4}$. Then a sufficient condition for the equilibrium to exists for any possible population dynamics is that, for each $i \in I$, $\hat{q}^i_t < \frac{{\alpha^i}^2\Delta V^i}{4} \leq 0.25$. Of course, even though this is a sufficient condition, it is quite restrictive since, for example, a group with q^i_t close to 1 must think to be less that a fourth of the overall population. For this reason, below in the paper we endogenize the conjectures to avoid this unreasonably large biases. Next proposition, however, uses this sufficient condition to characterize the dynamics for the cases in which conjectures do not change and the dynamic is defined at every $q^i_t \in [0,1]$. Define $Q^{ss} := \{q \in [0,1] : \dot{q}_t = 0\}$.

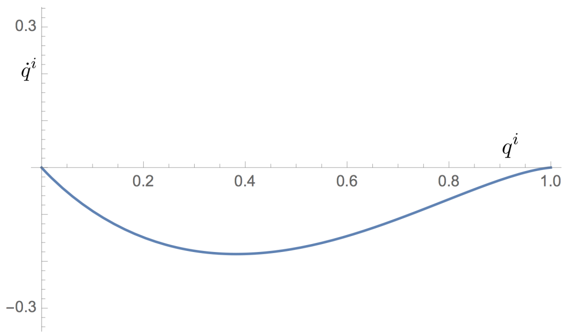
Proposition 9 *For each $i \in I$, let $\hat{q}^i \leq \frac{{\alpha^i}^2\Delta V^i}{4}$. $\{0,1\} \subseteq Q^i_{ss}$ and $q^i = 1$ or $q^i = 0$, or both, are stable. Moreover, if there exists a polymorphic steady state $q^* \in (0,1)$, $Q_{ss} = \{0, 1, q^*\}$, $q^i = 0$ and $q^i = 1$ are stable and $q^i = q^{*i}$ is unstable.*

Proof. in the Appendix B.2 □.

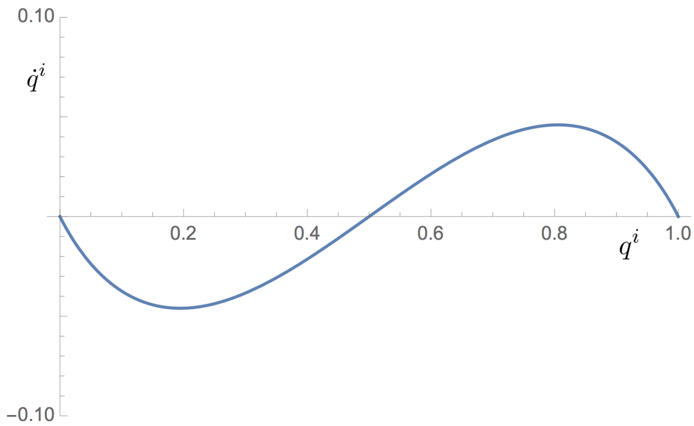
Figure 2.7 helps explaining the results of Porposition **??**. If conjectures are fixed and satisfy the sufficient condition for existence of confirmed conjectures at each point of the dynamics, there are three possibile cases. The first two cases, represented in panels (a) and (b) describe



(a)



(b)



(c)

Figure 2.7: Cultural dynamics $\dot{q}^i$ with $\alpha^i = \alpha^j = 1$: (a) $\hat{q}^i = 0.1$, $\hat{q}^j = 0.025$, $\Delta V^i = 0.5$, $\Delta V^j = 0.1$ (b) $\hat{q}^i = 0.025$ $\hat{q}^j = 0.1$, $\Delta V^i = 0.1$, $\Delta V^j = 0.5$ and (c) $\hat{q}^i = 0.1$ $\hat{q}^j = 0.1$, $\Delta V^i = 0.5$, $\Delta V^j = 0.5$.

the situation in which there exists only one homomorphic stable steady state. These two cases occur when the two groups have quite unbalanced conjectures. Consider for example panel (a). In this case group i has a conjecture about q^i that is four times the conjectures j have about q^j. Even if, through the feedback, parents update differently the conjectures about α^i and α^j, the large imbalance of the conjecture about population shares make i agents produce a much lower effort than j agents, at every point in the dynamics. Then j agents always increase in size and, at the end, win in the cultural dynamics. The same happens, with inverted roles, in panel j. Consider now panel (c) representing the case of quite balanced conjectures. In this case the difference in the exerted efforts is then given by the difference in the confirmed $\hat{\alpha}^i$ and $\hat{\alpha}^j$. Now, as we have seen in Figure 3.2, fixing $\hat{q}^i$, socialization effort increases with population shares. Indeed, one of the main results on our model is that it produces cultural complementarity with respect to actual population shares. In the case of panel (c), given that groups have same conjecutres about own share, then the group with the larger actual population share

produces the highest effort, and this determines the shape of the dynamics with one unstable polymorphic equilibrium.
If the sufficient condition for the existence of confirmed conjectures at any q is not satisfied, then we may have points in the dynamics in which the equilibrium does not exists and, thus, the dynamics cannot be determined. This is shown in Figure 8 and in which we show that, if conjectures of group i does not satisfy the condition, then they dynamics for small q^i cannot be determined. Notice that the condition implies that conjectures must be quite small. If this is not the case, then when q^i gets close to 0, conjectures are very far away from the real value, and then no $\hat{\alpha}^i$ can counterbalance the bias. The fact that the dynamics cannot be computed simply underlines the fact that in the long run conjectures must somehow be related, at least to a certain degree, to actual shares.

In the first graph we can observe that with both $\hat{q}^i = 0.12 < \bar{q}^i$ and $\hat{q}^j = 0.12 < \bar{q}^j$ the function is always continuous on the other hand if the conjectures of one or both types overcome that threshold the function became discontinuous, this mean that, given $q^i > \hat{q}^i$ for extrema value of q^i *selfconfirming* equilibria cannot exist.

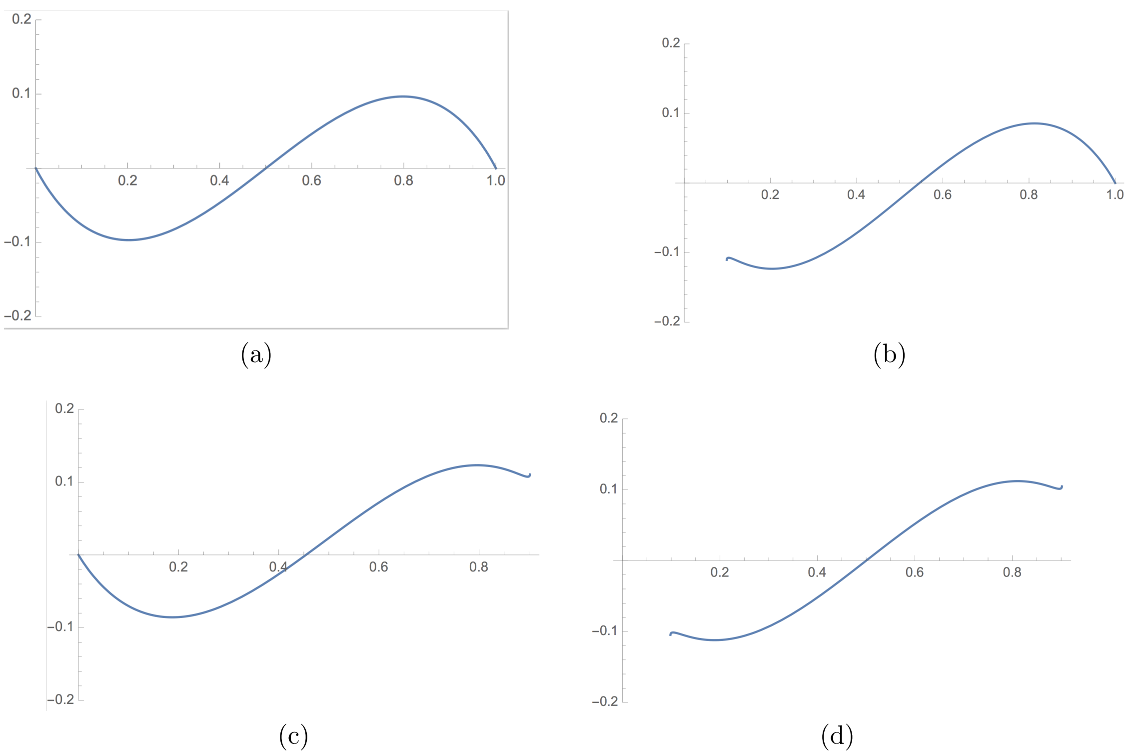


Figure 2.8: $\hat{q}^i = 0.12 < \bar{q}^i$ and $\hat{q}^j = 0.12 < \bar{q}^j$ (a) $\hat{q}^i = \hat{q}^i = 0.12$, (b) $\hat{q}^i = 0.12$ and $\hat{q}^j = 0.2$, (c) $\hat{q}^i = 0.2$ and $\hat{q}^j = 0.12$ and (d) $\hat{q}^i = \hat{q}^i = 0.2$

Proof of Proposition 9

It is trivial to see that $q^i = 0$ and $q^i = 1$ are steady states of $\dot{q}$.
In order to verify that may exists only an other possible steady state we should prove that the function ϕ has at most one solution. Moreover we have to show that if a solution exists the steady state associated with that solution is unstable. Therefore, we should prove that the function is monotone (uniqueness of solution) and increasing (instability of steady state, with cultural heterogeneity). Thus, we study the first derivative of ϕ with respsct q^i:

$$\frac{\partial \phi}{\partial q^i} = -\alpha^{i2}\Delta V^i - \alpha^{j2}\Delta V^j + \alpha^i \frac{\partial \xi^i}{\partial q^i} - \alpha^j \frac{\partial \xi^j}{\partial q^i}$$

where

$$\xi_i = \sqrt{4(q^i - \hat{q}^i)\Delta V^i + \alpha^{i2}(1-q^i)^2 \Delta V^{i2}}$$

$$\xi_j = \sqrt{4(1 - q^i - \hat{q}^j)\Delta V^j + \alpha^{j2} q^{i2} \Delta V^{j2}}$$

and

$$\frac{\partial \xi^i}{\partial q^i} = \frac{\Delta V^i (2 - \alpha^i \tau^i_{bv})}{\xi^i}$$

$$\frac{\partial \xi^j}{\partial q^i} = \frac{\Delta V^j (\alpha^j \tau^j_{bv} - 2)}{\xi^j}$$

$$\Rightarrow \frac{\partial \phi}{\partial q^i} = -\alpha^{i2}\Delta V^i - \alpha^{j2}\Delta V^j + \alpha^i \frac{\Delta V^i (2 - \alpha^i \tau^i_{bv})}{\xi^i} + \alpha^j \frac{\Delta V^j (2 - \alpha^j \tau^j_{bv})}{\xi^j}$$

it is decreasing both in $\Delta V^i, \Delta V^j, \alpha^i, \alpha^j$ and increasing in $\hat{q}^i$, $\hat{q}^j$ thus, to see if ϕ is increasing we can study the case $\Delta V^i = \Delta V^j = \alpha^i = \alpha^j = 1$ and $\hat{q}^i = \hat{q}^j = 0$

$$\frac{\partial \phi}{\partial q^i} = -2 + \frac{2-(1-q^i)}{\sqrt{4q^i + (1-q^i)^2}} + \frac{(2-q^i)}{\sqrt{4(1-q^i) + q^{i2}}} \geq 0$$

$$\frac{\partial \phi}{\partial q^i} = -2 + \frac{2-(1-q^i)}{\sqrt{2q^i + 1 + q^{i2}}} + \frac{(2-q^i)}{\sqrt{4 - 4q^i + q^{i2}}} \geq 0$$

$$\frac{\partial \phi}{\partial q^i} = -2 + \frac{2-(1-q^i)}{1+q^i} + \frac{(2-q^i)}{2-q^i} \geq 0$$

$$\frac{\partial \phi}{\partial q^i} = -2 + \frac{1+q^i}{1+q^i} + \frac{(2-q^i)}{2-q^i} = 0$$

Since the first derivative computed at his maximum is equal to zero then $\frac{\partial \phi}{\partial q^i} \geq 0$ for all parameters and ϕ it is always increasing monotone, therefore ϕ has no more than 1 solution.

$\Rightarrow \dot{q}$ has no more then three steady states: $q^i = 0, q^i = 1$ and $q^i : \phi = 0$.

Therefore it can be both positive and negative, if ϕ has 1 solution it is always negative and extrema steady states ($q = 1$ and $q = 0$) are stable, otherwise for very asymmetric parameters only one of the two estrema are stable ($q = 1$ or $q = 0$). Namely if an equilibrium with cultural heterogeneity does exist it is unstable.

□

Chapter 3

Non-Bayesian Social Learning and the Spread of Misinformation in Networks

Abstract

People are exposed to a constant flow of information about economic, social and political phenomena, nevertheless, misinformation is ubiquitous in the society. The paper studies the spread of misinformation in a social environment where agents receive new information each period and update their opinions taking into account both their experience and neighborhood's ones. We consider two sources of misinformation: permanent and temporary misinformation. The permanent one is modeled with the presence of stubborn agents in the network. Despite agents are exposed to constant flows of information, having stubborn in the network is enough to prevent the consensus, and thus the learning, to be reached. The distortion induced by stubborn agents in social learning depends on the "updating centrality", a novel centrality measure that identifies the key agents of a social learning process, and generalizes the Katz-Bonacich measure. Conversely, temporary misinformation, represented by shocks of rumors or fake news, has only short-run effects on the opinion dynamics. Results show that the consensus among agents is not always a sign of successful learning. Moreover, the consensus time is increasing with respect to the "bottleneckedness" of the underlying network, while the learning time is decreasing with respect to agents's reliance on their private signals.

Journal of Economic Literature Classification Numbers: D83, D85, D72, Z13
Keywords: Opinion Dynamics in Networks, Non-Bayesian Social Learning, Stubborn Agents, Speed of Convergence

3.1 Introduction

People form their beliefs and opinions about political, economic and social issues through the information they have. Nowadays, each person is exposed to a continuous stream of news about almost every subject. However, since no one has a direct access to the truth, and different pieces of information are dispersed among agents, people interact together and update their beliefs taking into account those of others. Moreover, agents tend to take in consideration others' belief to conform to their peers, or to some role models in society, even if they do not have better information. There is evidence, in fact, that people's opinions and decisions are affected by friends or neighbors or even influencers, such as sports celebrities, fashion bloggers, political leaders or commentators.[1]

In this scenario, the role of social media is of primary importance, nowadays, 62% of US adults use them as a source of news (Gottfried and Shearer, 2016). Social media, like Facebook and Twitter, allow agents to receive and share a lot of information in a very short time and to have an easy access to other's opinions. This, despite leading to a faster dissemination of news and faster social learning, leaves the door open to the spread of fake-news and misinformation or, in general, opinion manipulation (Del Vicario et al., 2016).

This raises important questions about the agents' learning process. Do social learning leads to a consensus among different individuals? Are agents able to effectively aggregate dispersed information about the underlying state of the world? How much room is there for belief manipulation and misinformation? Despite a large amount of data about economic, social and political phenomena disagreements are ubiquitous in society (Acemoglu and Ozdaglar, 2011). For example, people tend to disagree on many phenomena such as climate change, the effect of a flat tax or the guaranteed minimum income on the society, the effect of LGBT adoptions on the offsprings' nurture, or even the genuineness of the first moon landing, etc.. Therefore, we can deduce how consensus and learning are not always reached and there is room for indoctrination and the spread of misinformation. Moreover, since the consensus is a necessary but not sufficient condition for learning to take place, we wonder if social environments that guarantee a faster consensus would lead also to a faster learning.[2]

[1] For example, Coleman et al. (1966) discusses the role of doctors in prescription of new drugs, Reingen et al. (1984), Feick and Price (1987) and Godes and Mayzlin (2004) study the role of influencer marketing in brand choice by consumers, Martin and Bush (2000) asks if role models influence teenagers' purchase intentions and behavior, and Bush et al. (2004) analyzes the influence of sports celebrity on the behavioral intentions of a particular generation. Fainmesser and Galeotti (2018) propose a model of market interactions between influencers, followers, and marketers.

[2] It can be argued that does not exist a truth with respect to some of these examples and it is only a matter of preferences. However, in this paper, talking about consensus and learning, we refer to opinions about

The main contribution of this paper is to identify key players (i.e. nodes that if targeted are more effective in influencing the steady state opinion dynamics) and to analyze topological features that favor the spread of misinformation in a social learning framework where there is an underlying true state of the world and agents receive a constant flow of information. We study the spread of misinformation on a network composed by a set of agents, who each period receive noisy signals about the true state of the world and update their belief as a convex combination of the Bayesian posterior beliefs and a linear updating of neighborhood's beliefs, as in Jadbabaie et al. (2012). The analysis strongly depends on the assumption of normal distribution of beliefs, which is necessary to have a tractable close form solution. We consider two sources of misinformation, permanent and temporary.

One of the main tools used to control opinion dynamics in social media is the use of Social bots, algorithms that exhibit human-like behavior. They can be used to repeatedly share factious or even fake news and negatively (positively) comment to show a disagreement (or a consensus) higher than the true one.[3] Another source of permanent misinformation are "prominent agents" (media, firms or even politicians) that, systematically disseminate opinions and information to convey consent on themselves or on a particular idea that they support, an example are the so called "climate deniers", who promote skepticism about the scientific opinion on climate change (Oreskes, 2004). Both "prominent agents" and reliable social bots, that systematically spread misinformation, can be considered stubborn agents since they always support the same beliefs and do not learn neither by experience nor by peers, they act only to affect the outcome of others' social learning. Conversely, due to their use, many among social bots or troll are not credible and are unmasked in the long-run. Thus, they have only a short period effect in the opinion dynamics, through shocks of temporary fake news.

Thus, to model permanent misinformation, we assume the presence in the network of stubborn agents, who do not update their fixed beliefs and, in this way, are able to influence agents' social learning. Stubborn can represent "prominent agents", such as media, firms or politicians, who want to manipulate the opinion dynamics, social bots directly controlled from them, or

"objective" facts, that do not depend on personal or social preferences. E.g. we cannot say if a policy is "good" or "bad", since it depends on preferences, but, in principle, we can assess its objective effect in a particular economy.

[3]It was estimated (Varol et al., 2017) that the share of bots is between 9% and 15% of the total users active on Twitter, in the same way, many Facebook's users are "fakes". Social bots are used to spread fake news stories to influence political debates (Ratkiewicz et al., 2011), manipulate the stock market (Ferrara et al., 2016), and spread conspiracy theories (Bessi et al., 2015), among others. Moreover, Silverman (2016) shows that fake news stories are among the most shared on Facebook.

even agents who have incentives aligned with them.[4] Agents care about stubborn opinion either because they may believe that stubborn have information that they do not have or because they see stubborn as role models and thus have preferences to conform to them.

We show that having stubborn agents in the network is enough to prevent the consensus, and thus the learning, to be reached. Differently from classical naive learning model based only on linear updating (DeGroot, 1974; Golub and Jackson, 2010, among others) the steady state agents' opinions do not depend only on stubborn opinions and their centrality, but also on the true state of the world and on agents' reliance on their private signals (self-weights). The relevant centrality measure is a new one, the "updating centrality". The "updating centrality" identifies the key agent in a social learning process and, when agents are able to recall all their past signals, coincides with the Katz-Bonacich centrality of the network without self-weight. We further discuss the optimal strategy of a farsighted monopolistic stubborn who know the true state of the world and wants to minimize the distance between the steady state opinion vector and his own position. The stubborn face a quadratic lying cost increasing in the distance between her declaration and the truth. The main finding is a threshold value for the cost below which the stubborn declare an opinion more extreme than her true opinion.

To understand the effect of temporary misinformation, we study the speed of convergence of social learning in the nearby of steady state and without any stubborn agent, we focus only on the case in which agents are able to recall all their past signals. This analysis can be thought as the study of an exogenous shock which temporarily moves opinions away from the steady state; possible examples of this source of misinformation are the diffusion of unreliable fake news in social media or rumors in a social circle. Temporary misinformation does not affect opinion dynamics in the long-run, but have short period effects. Our results are based on spectral graph theory techniques. In particular, using Perron-Frobenius theorem and Cheeger's inequality we show, in line with previous literature, that the speed of reaching the consensus, inversely depends on the "bottleneckedness", and thus the homophily, of the underlying network. On the other hand, the speed of convergence toward the truth (speed of learning) mainly depends on the strength of weights that agents gives to their private signals and, surprisingly, it may not increase as you decrease the homophily, or in general the"bottleneckedness", of the network.[5] This is due to the fact that, if the level of private information is different across

[4]For example, influencer targeted by some firms or political party to support a particular product or policy. Among theoretical results, Galeotti and Goyal (2009) provides a model about strategic diffusion through influencer, while Fainmesser and Galeotti (2015) and Fainmesser and Galeotti (2016) studies influencer marketing in both monopolistic and oligopolistic framework.

[5]Surprisingly with respect the results of Golub and Jackson (2012).

agents, the learning of better informed agents is slowed down by others and this reduce the speed of learning of the whole society

The article is organized as follows. In Section 3.1.1 we offer a brief literature review. Section 3.2 lays out the formal framework of the model where we present, and microfound, agents' updating rule. Section 3.3 studies the effect of permanent misinformation introducing the presence of stubborn agents in the network. Section 3.3.1 characterizes the steady state opinions' vector and define the "updating centrality". Section 3.3.2 is devoted to comparative statics and in particular to measure the marginal distortion due to an outgoing link from a stubborn. Section 3.3.3 studies the optimal opinion to declare for a monopolistic stubborn who want to affect the opinion dynamics. Section 3.4 studies temporary misinformation discussing the speed of convergence toward consensus and the speed of learning, showing when they coincide or differ and why. Section 3.5 concludes.

3.1.1 Literature Review

The main purpose of this project is to create a deeper link between the literature on learning in networks and the literature on the optimal targeting of individuals to diffuse (mis)information or opinions in a social network. Moreover, we show how a constant flow of information may seriously change the steady states opinions' vector and the convergence time with respect to standard DeGroot based models, where agents receive at most one signals at the first period. This paper refers to different streams of literature.

Opinion Dynamics and Learning in Networks The literature about opinion dynamics and learning on networks can be divided into two main approaches: *Bayesian* and *non-Bayesian* learning models. (Golub and Sadler, 2016, for a survey). In particular, the social learning process of this paper belongs to the non-Bayesian stream of literature, which began, and strongly relies, on the famous DeGroot model, DeGroot (1974).[6] Starting from the standard DeGroot model, DeMarzo et al. (2003) make explicit the role of the network, while Golub and Jackson (2010) derive general conditions on the adjacency matrix to ensure the reaching of consensus. Concerning our purpose, the main limit of DeGroot like models is that agents can receive signals about the true state of the world only at the first period and then they update their beliefs aggregating others beliefs, as in Golub and Jackson (2010). In such a cases is very easy to manipulate others' opinions and spread misinformation in the network. Therefore, to better analyze the spread of misinformation we consider agents who receive signals about the

[6]Some paper belonging to the Bayesian learning literature are Bala and Goyal (1998), Gale and Kariv (2003), Rosenberg et al. (2009), Acemoglu et al. (2011), Mossel et al. (2015). For a deeper discussion about both Bayesian and non-Bayesian paradigms we remand to the survey Golub and Sadler (2016).

true state of the world at each period. As social learning process, we use the one proposed by Jadbabaie et al. (2012), where agents combine their personal experience and the views of their neighbors as a convex combination of the Bayesian posterior beliefs, given their personal signals, and a linear updating of neighborhood's beliefs.

Behavioral explanation for the Learning Process A contribution of our paper is to use a beauty-contest like utility function (in the spirit of Morris and Shin (2002)) to propose an explanation for the mechanisms behind the specific social learning process used (Jadbabaie et al., 2012). Beauty-contest like utility functions are widely used in the literature about learning and opinion dynamics. For example, Bindel et al. (2015) use a similar utility function show the correspondence with the DeGroot model and compute its the inefficiency (with respect to other non-equilibrium strategies) through the price of anarchy. Buechel et al. (2015) study the opinion dynamics when agents may misrepresent their own opinion by conforming or counter-conforming with their neighbors. In Olcina et al. (2017) the beauty-contest like utility function is used to study the norms' assimilation of ethnic minorities. A similar payoff structure is used also in Bolletta and Pin (2018) where a dynamic process, with co-evolution of both individual opinions and network, is characterized. Molavi et al. (2018), studies the behavioral foundations of non-Bayesian models of learning over social networks under the main behavioral assumption of "imperfect recall" of others' beliefs, showing that, social learning rules have a log-linear form, as long as imperfect recall is the only point of departure from Bayesian rationality. In the end, Dasaratha et al. (2018) study a set of Bayesian agents that learn about a moving target, the main result is that, under incomplete information, a fully Bayesian learning model can be tractable as the standard DeGroot heuristic.

Stubborn Agent As channel of permanent misinformation, we make use of stubborn agents. Yildiz et al. (2013) use the concept of stubborn agents, firstly proposed by Mobilia (2003) and Mobilia et al. (2007),[7] in a binary opinion dynamics framework. While, models as Grabisch et al. (2017) and Mandel and Venel (2017) study influences and targeting in networks, through stubborn agents, using the DeGroot model for the learning process.

Targeting in Networks Since the position of the stubborn in the network is of primary importance, results refer also to the literature that studies the problem of optimal targeting in networks (Bloch, 2016, for a survey). Papers as Galeotti and Goyal (2009), Candogan et al. (2012) and Fainmesser and Galeotti (2015) study targeting and pricing problem from a monopolistic point of view, while Goyal et al. (2014), Fainmesser and Galeotti (2016) and Bimpikis et al. (2016) deal with the same problem in the competitive case. Our main contribution is

[7]They use the term zealot instead of stubborn.

to make the first attempt to consider a targeting problem, through stubborn, in a society where agents receive each period signals about the true state of the world, therefore the tension between the learning and the targeting force is very relevant in our problem.

Key Player Important contributions, in economics, to targeting problems stem also from the literature that aims to identify the key agent in a network (Zenou, 2016, for a survey). Among the first and most important results there is Ballester et al. (2006), which define the "intercentrality" measure, a centrality measure that takes into account not only a player's centrality but also her contribution to others' centrality. While Banerjee et al. (2013) derive a measure of "diffusion centrality" that discriminates between information passing and endorsement. We contribute to this literature deriving the concept of "updating centrality" measure, which directly depends on the agent's (linear) updating rule.

Speed of Learning Since the effect of temporary misinformation has only short run effect, we analyze the speed of learning and the convergence to consensus. Golub and Jackson (2012), among others, find the most striking result in this literature, the authors examine how the speed of learning of average-based updating processes (as DeGroot) depends on homophily, showing that convergence to a consensus, is slowed by the presence of homophily.[8] Our contribution is to show that, if agents receive signals at each period, that anchor them to the truth, then the speed learning can be different from the speed of convergence and is not necessarily slowed by the presence of homophily.

3.2 The Model

In this Section we introduce the baseline model, which strongly relies on Jadbabaie et al. (2012). In Section 3.3 and 3.4 we study the effect of systematic and temporary misinformation on agents' social learning.

The Society The society is represented by a graph $\mathcal{G}(N, \boldsymbol{\mathcal{A}})$, where $N = \{1, 2, ..., n\}$ is the set of finite nodes or agents and $\boldsymbol{\mathcal{A}} \in [0, 1]^{N \times N}$ is the matrix that captures the interaction patterns among agents in N. In particular, a_{ij} is the ij-th entry and represents the weight that i gives to agent j, namely how much i listen j in proportion to others agents in N. Each agent $i \in N$ divides her attention between himself and the other agents in N, thus the matrix is row-stochastic so that its entries across each row are normalized, $\sum_{j \in N} a_{ij} = 1$. We keep

[8]In Golub and Jackson (2012), the speed of convergence to a consensus is equivalent to the speed of learning.

the analysis as general as possible considering a directed network where the interactions can be asymmetric and one-side, so that $a_{ij} >$ while $a_{ji} = 0$.[9]

States of Nature and Signals The finite set of possible states of nature is $\Theta \subseteq \mathbb{R}$ where θ^* is the true state of world. Conditional on the true state of the world θ^*, at each period agents observe noisy signals $\omega_t = (\omega_{1,t}, \omega_{2,t} ..., \omega_{n,t})$. Signals, that a generic i receive during her life, are drawn from a distribution with mean $E[\omega_i] = \theta^*$, variance $\sigma^{2\omega_i} > 0$, and precision $\tau^{\omega_i} = \frac{1}{\sigma^{2\omega_i}}$, for all $i \in N$. We further assume that for each agent signals are *i.i.d.* over time and agents do not need to have any information about signals generation processes.

Agents' Opinion Each agent i, has at each time t a normal probability distribution (or probabilistic belief) over the possible state of the world $p_{i,t}(\theta) \in \Delta\Theta$.[10] The opinion (belief) of i, at each t, is $\mu_{i,t} = \int_\Theta \theta p_{i,t}(\theta)d\theta$, the first moment of her probability distribution over Θ. While, $\sigma^{2p}_{i,t}$ and $\tau^p_{i,t}$ are respectively the variance and the precision of the probability distribution of i over the possible states at t.

Social Learning We assume that agents observe through communication their neighbors' beliefs. The observed beliefs are used, jointly with private signals received at each period, to update beliefs about the underlying state of the world. The matrix $\boldsymbol{\mathcal{A}}$ describe weights of communication and reliance on their private signals.[11] In this model, agents have always "imperfect recall" about others' beliefs, namely, treat them as sufficient statistics for the entire history of their observations.[12] We do not impose any restriction about the recall of past private signals studying both "imperfect" and "perfect recall". With "imperfect recall", agents takes into account only the last signal received in the updating process. On the other hand with "perfect recall" agents are able to recall and use the whole history of received signals.

Formally, the updating rule of probabilistic belief for each agent $i \in N$ is assumed to be a convex combination of the Bayesian updating β, and the average probabilistic beliefs of their neighborhood

[9]We can trivially restrict the analysis to directed networks, where $a_{ij} = a_{ji}$, results do not change.

[10]Notice that the assumption of normal belief is useful to have in next section tractable (linear) closed form solution.

[11]Notice that we are assuming that individuals update their true opinions in a non-fully rational way. In fact, if individuals were fully rational, they would perfectly account for repetition of information. Empirical evidence strongly suggests that individuals are not fully-rational in these settings. For example, laboratory experiments shows that in both complex networks (Grimm and Mengel, 2014) and also in small social networks with common knowledge (Corazzini et al., 2012), people fail to properly account for repetitions of information.

[12]We remand to Molavi et al. (2018) for a deeper discussion about the implication of "imperfect recall" of others' belief in social learning.

$$p_{i,t+1}(\theta) = a_{ii}\beta_{i,t+1} + \sum_{j \in \mathcal{N}_i} a_{ij} p_{j,t}(\theta) \tag{3.1}$$

Where

$$\beta_{i,t+1} = \frac{l(\omega_{i,t+1}|\theta)p_{i,t}(\theta)}{\int_\Theta l(\omega_{i,t+1}\theta)p_{i,t}(\theta)d\theta}$$

is the Bayesian posterior belief at t for agent i and $l(.|\theta^*)$ is the likelihood function that generates signals ω. Elements a_{ii} are the self-reliance of each agent i, while a_{ij} represents how agent i weights j's beliefs.

To study opinion dynamics, in this paper, we focus on the mean of the distribution. At each time t the belief (opinion) of agent i is the first moment of the probability distribution (3.1), namely

$$\mu_{i,t+1} = a_{ii} \int_\Theta \theta \beta_{i,t+1} d\theta + \sum_{j \in \mathcal{N}_i} a_{ij} \mu_{j,t}. \tag{3.2}$$

Before to discuss, in next session, permanent sources of misinformation it is important to provide a behavioral explanation for our particular social learning rule.

Microfoundation There can be many reasons for which agent may aggregate others' beliefs. For example, since agents do not have complete information about the distribution from which signals are drawn, they want to aggregate others' information. Another possibility is that they have preferences to conform to others agents or to specific role models in the society. These considerations lead us to offer a micro-foundation for the updating rule (3.1) of agents in the society with the following utility function.[13]

$$u_{i,t}\left(p_{i,t}(\theta), p_{-i,t}(\theta)\right) = -\left(p_{i,t}(\theta) - 2a_{ii}\beta_{i,t+1}\right)^2 - \left(p_{i,t}(\theta) - 2\sum_{j \in \mathcal{N}_i} a_{ij} p_{j,t}(\theta)\right)^2 \tag{3.3}$$

Solving first order conditions for (3.3) we find exactly the updating rule in (3.1).[14]

In next section we introduce stubborn agents in the network. Agents may care about opinions of stubborn for the same reasons.

[13] The first appearance of a similar utility function in economics is due to Morris and Shin (2002). We can find it in similar frameworks in Bindel et al. (2015), Olcina et al. (2017), Dasaratha et al. (2018), Bolletta and Pin (2018), among others.

[14] Which is equivalent to Jadbabaie et al. (2012).

3.3 Permanent Misinformation

In this section, we introduce in the society a set of stubborn agents S. Stubborn represent "prominent agents" who have a fixed opinion and repeatedly share factious information or actual misinformation in social media (or social cliques). A generic stubborn $s_s \in S$, is a particular agent that is not affected by the opinion of others and never revises her opinion θ_{s_s}, namely $a_{s_s s_s} = 1$ and $a_{s_s i} = 0$ for all $i \in N$.[15]

As discussed in the previous section, agents may care about stubborn's opinions, in (3.3), for many reasons. For example, agents can ignore the presence of stubborn in the society, or they may believe that stubborn have information that they do not have or they see stubborn as role models and thus have preferences to conform to them. Thus, equation (3.3) still holds, an agent i takes into consideration the opinion of stubborn s_s if $s_s \in \mathcal{N}_i$.

The new adjacency matrix $\boldsymbol{\mathcal{A}}_s$ of the society with stubborn, is defined as[16]

$$\boldsymbol{\mathcal{A}}_s = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n1} & a_{1s1} & a_{1s2} \\ a_{21} & a_{22} & \dots & a_{2n} & a_{2s1} & a_{s2} \\ \dots & \dots & & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 & 1 & 0 \\ 0 & 0 & \dots & 0 & 0 & 1 \end{bmatrix} = \left[\begin{array}{c|cc} \boldsymbol{\mathcal{A}} & \boldsymbol{a}_{s1}, & \boldsymbol{a}_{s2} \\ \hline \mathbf{0} & 1 & 0 \\ \mathbf{0} & 0 & 1 \end{array}\right]$$

Where $\boldsymbol{a}_{s1}$ and $\boldsymbol{a}_{s2}$ are the column vectors composed by the weight that each agent i gives to the stubborn. Notice that, with stubborn agents in the society, $\boldsymbol{\mathcal{A}}$ is row-substochastic and $\boldsymbol{\mathcal{A}}_s$ is row-stochastic, such that $\sum_{j\in N} a_{ij} + \sum_{s_s\in S} a_{is_s} = 1$, for each $i \in N$.

Stubborn receive always the same signals ω_{s_s} for all $s_s \in S$ with $E[\omega_{s_s}] = \theta_{s_s}$ and zero variances. Since stubborn agents never revise their opinions, neither through social interaction nor through signals, their beliefs are fixed over time, $\mu_{s_s,t+1}(\theta^*) = \mu_{s_s,t}(\theta^*) = \theta_s$.[17]

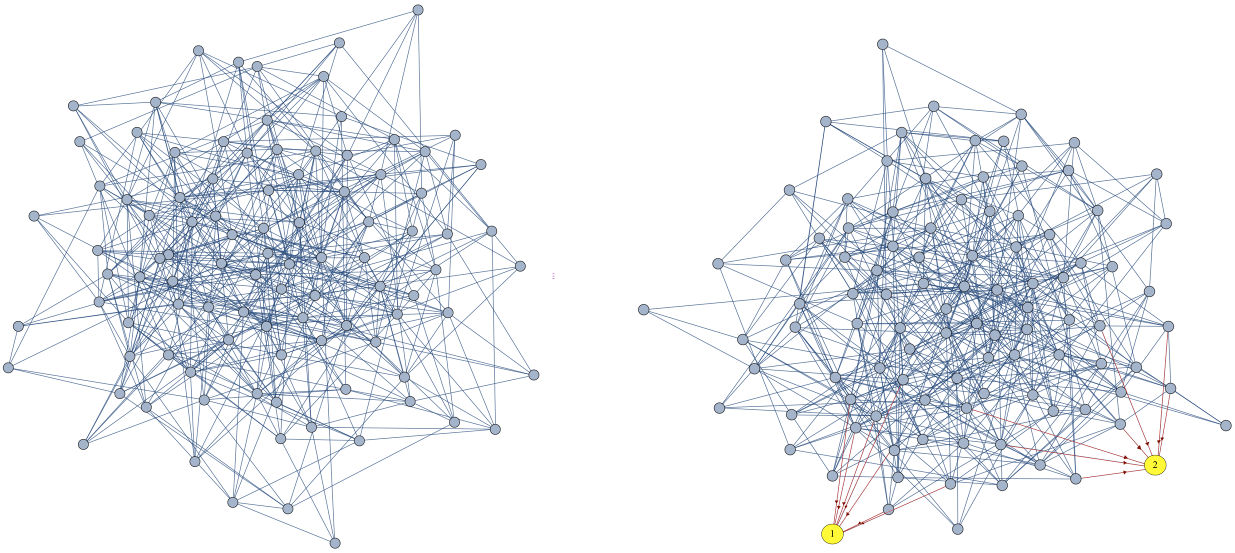
With stubborns, the updating of non-stubborn agents is

$$\mu_{i,t+1} = a_{ii} \int_{\Theta} \theta \beta_{i,t+1} d\theta + \sum_{j\in N} a_{ij}\mu_{j,t} + \sum_{s_s\in S} a_{is_s}\theta_{s_s} \tag{3.4}$$

[15] Our stubborn agent are the same as the stubborn in Yildiz et al. (2013) and the zealot in Mobilia (2003) and Mobilia et al. (2007).

[16] Here we consider the case with only two stubborn agents, but we easily can generalize it to any number of stubborn.

[17] We can provide a micro-foundation for stubborn, but in this case it is trivial.



(a) (b)

Figure 3.1: (a) Network with 198 nodes and 486 edges $\mathcal{G}(N, \boldsymbol{\mathcal{A}})$ (b) same network with 2 stubborn agents with 7 link each $\mathcal{G}'(N \cup S, \boldsymbol{\mathcal{A}}_s)$

Where the weight that i gives to a generic stubborn ($s_s \in S$) is $a_{is_s} \neq 0$ if and only if $s_s \in \mathcal{N}_i$. [18] Therefore, (3.2) and (3.4) have exactly the same meaning when $s_s \in \mathcal{N}_i$.

In next section, 3.3.1, we discuss the effect of stubborn opinions on the steady state opinions' vector of other agents, we consider the case with only two stubborn agents, but results hold for any number of stubborn.

3.3.1 Characterization of Steady State Opinions' Vector

We now write the opinions' updating rule (3.4), for the whole society, in matrix form. We decompose the adjacency matrix $\boldsymbol{\mathcal{A}}$ as the sum of the diagonal matrix containing all the weight that agents give to their private signals (self-loops), $\boldsymbol{D} = diag[a_{11}, ...a_{nn}]$, and the adjacency matrix of the same network without self-loops, $\boldsymbol{A} = \boldsymbol{\mathcal{A}} - \boldsymbol{D}$. In particular, $\boldsymbol{D}$ represents the Bayesian part of agents' updating and $\boldsymbol{\mathcal{A}}$ the linear (DeGroot) one. We further define $\bar{\beta}_{i,t} = \int_\Theta \theta \beta_{i,t+1} d\theta$ as the first moment of the Bayesian posterior, and $\bar{\boldsymbol{\beta}}_{t+1}$ as the column vector containing the all $\bar{\beta}_{i,t}$. $\boldsymbol{a}_{s1}$ and $\boldsymbol{a}_{s2}$ are vectors containing all influences of stubborn over agents.

[18]Notice that agents may both not to be aware of the presence of stubborn in the network and can consider them as other agents, or even they may be aware of their presence but they can consider them for other sociological reasons.

In matrix form the updating of non-stubborn agent is

$$\boldsymbol{\mu}_{t+1} = \boldsymbol{D} \cdot \bar{\boldsymbol{\beta}}_{t+1} + \boldsymbol{A}\boldsymbol{\mu}_t + \boldsymbol{a}_{s1}\theta_{s_1} + \boldsymbol{a}_{s2}\theta_{s_2}. \tag{3.5}$$

Before to characterize the steady state of opinion dynamics, let us define $\boldsymbol{G}$ as the diagonal matrix with $\gamma_i = \frac{\tau_i^p}{\tau_i^p + \tau^\omega}$ as entries, where τ_i^p is the precision of the probability distribution for i at the steady state.

Proposition 1 *If, for all agents, the probability distribution over* Θ *at* $t = 0$ *(the prior) is normally distributed and "imperfect recall" of past signals holds, then steady state vector of opinions (beliefs) is*

$$\boldsymbol{\mu} = \boldsymbol{C}\left(\boldsymbol{D}(\boldsymbol{I} - \boldsymbol{G})\boldsymbol{\theta}^* + \boldsymbol{a}_{s1}\theta_{s1} + \boldsymbol{a}_{s2}\theta_{s2}\right), \tag{3.6}$$

where

$$\boldsymbol{C} \cdot \boldsymbol{1} = (\boldsymbol{I} - \boldsymbol{DG} - \boldsymbol{A})^{-1} \cdot \boldsymbol{1} \tag{3.7}$$

is the vector of "updating centrality".

Proof. In the Appendix □.

Proposition 1 shows that the steady state opinion vector is a linear combination of the underlying state of the world and stubborn opinions. Moreover, the relative influence that an agent has over opinions' steady state, and thus the weights of the linear combination, depends on the "updating centrality" and on the weight that she gives to her private signals, and to stubborn agents, respectively. The "updating centrality", $\boldsymbol{C} \cdot \boldsymbol{1}$, represents the relative influence of each node at the steady state of a particular social learning framework.[19] This result is particularly relevant with respect to the literature about targeting and the key player in networks (Bloch, 2016; Zenou, 2016, for reviews of these literature). If a firm (or political party) wants to target agents who receive a constant flow of information to disseminate their message, it should target the more central ones with respect the "updating centrality" measure, previously defined. Differently, from other network centrality, the "updating centrality" of an agent, does not depends only on the topology of the underlying network, but positively depends on the steady state belief's precision τ_i^p. Namely, according to this centrality measure, agents that end up to have less sharp opinions at the steady state, ceteris paribus, are less central than others. The intuition is straightforward, the more the position of an agent allows her to have

[19] Notice that the vector of "updating centrality" depends on the updating rules, and priors' distribution. We are going to characterize the "updating centrality" only in the specific case of our model, but, for example, it is trivial to see that for the DeGroot like models the "updating centrality" is nothing but the eigenvector centrality.

a sharper opinion at the steady state (higher precision), the more, if targeted, her opinion remains close to the stubborn opinion, and in this way the targeting action results to be more effective. Roughly speaking the steady state belief's precision τ_i^p a measure of how effectively stubborn can convey her opinion to and through the targeted agent i.

As a particular case we consider the situation in which agents are more rational and are able to recall all past signals.[20]

Corollary 1.1 *If, for all agents, the probability distribution over Θ at $t = 0$ (the prior) is normally distributed and 'perfect recall" of past signals holds, then steady state vector of opinions (beliefs) is*

$$\boldsymbol{\mu} = \boldsymbol{C}\left(\boldsymbol{D}\boldsymbol{\theta}^{*} + \boldsymbol{a}_{s1}\theta_{s1} + \boldsymbol{a}_{s2}\theta_{s2}\right), \tag{3.8}$$

where

$$\boldsymbol{C}\cdot\mathbf{1} = (\boldsymbol{I}-\boldsymbol{A})^{-1}\cdot\mathbf{1} \tag{3.9}$$

is the vector of "updating centrality".

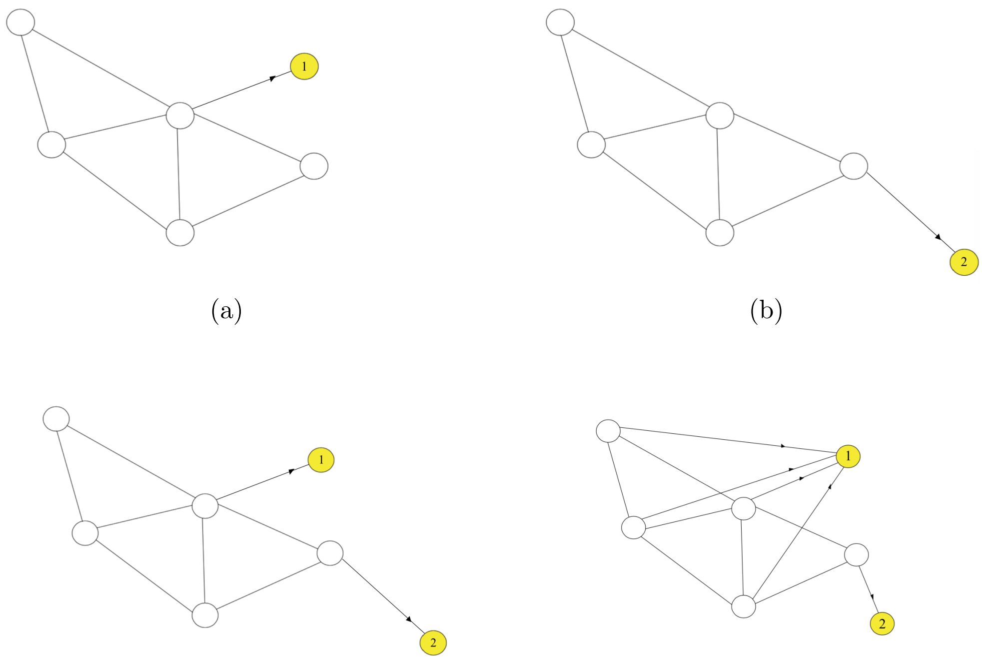
Proof. In the Appendix □.

Proposition 1 and Corollary 1.1 show that the presence of stubborn agents in the network prevents the consensus to be reached at steady state. Under "perfect recall" the probability distributions over the possible state of the world is degenerate with three picks: the true state of the world, opinion of stubborn 1 and opinion of stubborn 2. The relationship with the "updating centrality" holds even when agents are able to recall past signals. In this case, since the belief's precision τ_i^p is the same for all agents, the "updating centrality" depends only on the underlying network and coincides with the Katz-Bonachic centrality of the network without self-weight (with parameter 1).

Example 1 Consider a society of five agent, $N = \{1, 2, 3, 4, 5\}$ and two stubborn $s = \{s_1, s_2\}$, with the true state being $\theta^* = 10$ and stubborn's opinions , $\theta_{s_1} = 15$ and $\theta_{s_2} = 5$, respectively. Let assume that all agents in N have positive self-reliance and perfect recall of past signals. As for the social structure, we consider the four possibile situations described in Figure 2. We assume that the network is a normalized 0,1 network, where the intensity of each link is $\frac{1}{|\mathcal{N}_i|}$, with $i \in \mathcal{N}_i$, for all $i \in N$. Thus, since each node has a positive self reliance, if (for example) a node has 3 links the intensity of each link and the self reliance are both $\frac{1}{4}$. Using results

[20] I disregard intermediate cases, where agents recall some of the past signals, since they do not add much to the understanding of social learning.

from Corollary 1.1, it is trivial to compute the steady state opinion vector and the distortion created by each link toward a stubborn. In order to see the total effect of one or two stubborn in different social structures we can compute the average steady state opinion $\boldsymbol{\mu}'\mathbf{1}/n$ which results to be 10.75 in (a) 9.3101 in (b) 10.0609 in (c) and 11.5000 in (d). These example show how the effect of stubborn agents on the opinion vector in steady states directly depends on the centrality of targeted agents, the more the central is the agents connected with the stubborn the higher is stubborn centrality and thus his influence. ■



(a) (b) (c) (d)

Figure 3.2: Networks composed by agents with "perfect recall" and the intensity of each link of i is $\frac{1}{|\mathcal{N}_i|}$ used in Example 1 (self-loop is omitted in figures) and Example 2 (no self-loops at all). $\theta^* = 10$, $\theta_{s1} = 15$, $\theta_{s2} = 5$ the average opinions $\boldsymbol{\mu}'\mathbf{1}/n$ in the two examples are: (a) 10.75, (b) 9.3101, (c) 10.0609, (d) 11.5000 (in Example 1) and (a) 15 (b), (c) 9.8455, (d) 12.7500 (Example 2).

As we have seen in Proposition 1 and Corollary 1.1, the underlying state of the word θ^* plays a role in determining the steady state opinions' vector $\boldsymbol{\mu}$. To better understand how the constant flow of new information each period is important to agents' learning we consider, in the next example, a set of agents that do not listen the received signals, thus update their belief as in a standard DeGroot framework.[21]

[21]This updating rule is the same used in Golub and Jackson (2010, 2012), in such a case our model with stubborn is comparable to models as Yildiz et al. (2013) or Grabisch et al. (2017).

Example 2 Let us consider the same society with the same state of the world and stubborn as in *Example 1*, but let all elements a_{ii} in $\boldsymbol{D}$ be equal to zero. Namely, agents update their beliefs without taking in account the signals received at each period, thus the updating rule is exactly the standard DeGroot linear updating model. Notice that, in this case $i \notin \mathcal{N}_i$ and thus if (for example) a node has 3 links the intensity of each link is $\frac{1}{3}$. We can easily verify how the average steady state opinion is 15 in (a) and 5 in (b), the variance is zero in both case. Namely, if there is only one stubborn in a network where agents linearly update their beliefs without constant signals (as in DeGroot models) the opinion of each agent always converges to the stubborn's opinion. On the other hand, if there is more than one stubborn in the network, as in (c) and (d), the steady state depends only on the centrality of agents connected with stubborn and the stubborn's opinion. In particular, $\boldsymbol{\mu}'\mathbf{1}/n$ is 9.8455 in (c) and 12.7500 in (d). Notice that, in *Example 1*, where agents took into consideration the received signals ($a_{ii} \neq 0$), the Bayesian updating based on a constant flow of signals mitigates the stubborn influence on the steady state opinion vectors. ■

Corollary 1.1 tells us that, if agents are able to recall past signal their opinions are closer to the truth than with "imperfect recall", nevertheless having 'perfect recall" of signals is not enough to have the convergence toward the truth. However, there exist particular situations in which a society with stubborn reaches the truth. It is trivial to see that all elements in $\boldsymbol{\mu}$ are equal to the truth θ^* if agents do not give importance to the stubborn, $\boldsymbol{a}_{s1} = \boldsymbol{a}_{s2} = 0$, or if the stubborns' opinion is equal to the truth itself, $\theta_{s1} = \theta_{s2} = \theta^*$. Moreover, if stubborn have the same influence over the society, namely if $\boldsymbol{a}_{s_1} = \boldsymbol{a}_{s_2}$, then the society can converge to the truth whenever $\theta_{s_1} = 2\theta^* - \theta_{s_2}$. Notice that in these two case the welfare is maximized both from the policy maker's and agents' prospective. We remand to the Appendix B for details and a deeper discussion about approaching the truth where there are stubborn agents in the society.

We can see that having more than one stubborn agent may facilitate learning if stubborn are evenly distributed around the truth. Therefore, from a policy maker point of view, it may make sense to facilitate the entry of stubborn agents, with different opinions, once there is already one in the network. This shows us that, for example, if the presence of a factious social media is recognized in the society, having more social media with different opinion may be desirable. Notice that this argument is valid only if the network is exogenous, as in this paper. On the contrary, if agents have the possibility to choose their connection then, depending on the network structure and the strength of the signals, stubborn can be isolated in the long-run, or there can arise different isolated communities where agents as opinion very close to the stubborn.

In next two section (3.3.2 3.3.3), to convey results and, at the same time, to maintain formal simplicity, we consider a network with only one stubborn agent. We study, in 3.3.2, the distortion induced by a stubborn and, in 3.3.3, the optimal declaration of a sophisticated stubborn.

3.3.2 Marginal Distortion Induced by Stubborn

In order to better understand the effect of stubborn agents on the social learning, we study the effect of increasing (or decreasing) the influence of a stubborn agent s over a generic agent i of α_{is}.[22]

Since the listening matrix should remain normalized (row-stochastic) if we define $\hat{a}_{is} = a_{is} + \alpha_{is1}$ the new influence of a generic stubborn s on agent i then $\sum_{j \neq \{i,s\}} \hat{a}_{ij} = \sum_{j \neq \{i,s\}} \hat{a}_{ij} - \alpha_{is}$ ($\sum_{j \neq \{i,s\}} \alpha_{ij} = -\alpha_{is}$ and therefore $\alpha_{is} = -\sum_j \alpha_{ij}$) we further assume that a_{ii} remain fixed for all i, that $\alpha_{is} = \alpha$ and $\alpha_{ij} = \frac{\alpha}{n}$, for simplicity.

Let $\hat{\boldsymbol{\mathcal{A}}} = \boldsymbol{\mathcal{A}} - \mathbf{e}_i(\mathbf{g}_i)'$ be the modified matrix where $\mathbf{g}_i$ is the n-dimensional listening column vector that has: (i) 0 in its ith position, (ii) $\alpha/|\mathcal{N}_i|$ in all jth positions different from the one associated with the stubborn; $\mathbf{e}_i'$ is the is the n-dimensional row vector that has a 1 in its ith position and 0 elsewhere.

The next proposition describes the effect of introducing one stubborn agent in the society composed only by non-stubborn agents.

Proposition 2 *In a society without stubborn agents, if one agent create a link of intensity α with a stubborn, the marginal effect is*

$$\Delta \boldsymbol{\mu} = (\boldsymbol{C} - \boldsymbol{X})\alpha \boldsymbol{e}_i \theta_s - \boldsymbol{X}\boldsymbol{D}(\boldsymbol{I} - \boldsymbol{G})\boldsymbol{\theta}^* \tag{3.10}$$

and the steady state opinion vector is

$$\hat{\boldsymbol{\mu}} = (\boldsymbol{C} - \boldsymbol{X})\left(\boldsymbol{D}(\boldsymbol{I} - \boldsymbol{G})\boldsymbol{\theta}^* + \alpha \boldsymbol{e}_i \theta_s\right) \tag{3.11}$$

Moreover,

$$\boldsymbol{X} \cdot \mathbf{1} = \frac{\boldsymbol{C}\boldsymbol{e}_i(\boldsymbol{g}_i)'\boldsymbol{C}}{1 + (\boldsymbol{g}_i)'\boldsymbol{C}\boldsymbol{e}_i} \cdot \mathbf{1}$$

describe the distortion induced by the stubborn on agents' "updating centrality".

[22] Since in this section and in the next one we consider only one stubborn we call it s, instead of s_s.

Proof. In the Appendix □.

The distortion induced by a stubborn in the opinion steady state vector is $|\Delta\boldsymbol{\mu}|$. From (3.10), it is evident that the distortion is increasing in the distance between θ_s and θ^*. Moreover, as shown by the distortion term $\boldsymbol{X}$, an increase (decrease) in stubborn influence creates a distortion in the centrality of all agents in the network, the new "updating centrality" is $(\boldsymbol{C} - \boldsymbol{X})$. [23]

If we consider a benevolent policy maker who want to minimize the distance between agents' opinion and the truth, (for example with utility function $u_p(\boldsymbol{\mu}) := -(\boldsymbol{\mu} - \boldsymbol{\theta}^*)^2$, the distortion $|\Delta\boldsymbol{\mu}|$ represents a measure of the policy maker's welfare loss. The welfare of policy maker is maximum when all the agents learn the truth, $|\Delta\boldsymbol{\mu}|$.

3.3.3 Monopolistic Stubborn Agent Problem

We have, until now, considered naive stubborn agents that always declare exactly the opinion that they want to disseminate. We now consider a sophisticated stubborn that is able to optimally choose the opinion to declare in order to maximize the diffusion of the opinion that she really want to disseminate.

The stubborn knows the true state of the world θ^* and it is farsighted, namely she is able to compute the steady state opinion vector $\boldsymbol{\mu}$. The stubborn, given a fixed influence over agents, can declare any opinion $\theta_s^d \in \Theta$ such that minimize the distance between agents' opinion and her own true opinion θ_s. We further assume that the stubborn faces a cost of lying (Kartik et al., 2007; Kartik, 2009). The cost of lying is assumed to be quadratic and proportional to difference between the true state of the world θ^* and the declared opinion θ_s^d, and is parameterized by $k \in \mathbb{R}_+$, the intensity of the lying cost.

The interpretation of the intensity of the lying cost, k, can be manifold, it can be thought as the punishment (fines) for having spread fake news, or as a cost to convince others about the reliability of your opinion, or can even represent the expected loss in credibility due to a too extreme declaration.

[23] In this section, we study to a situation in which stubborn agents are not connected with the rest of the society ($\boldsymbol{a}_{s1} = \boldsymbol{a}_{s2} = \boldsymbol{0}$). In the Appendix we extend the results to generic $\boldsymbol{a}_{s1}$ and $\boldsymbol{a}_{s2}$.

The stubborn choses to declare θ_s^d that solve the following problem

$$\max_{\theta_s^d} \; u_s(\boldsymbol{\mu}) = \; -(\boldsymbol{\mu} - \mathbf{1}\theta_s)^2 - k(\theta^* - \theta_s^d)^2 \tag{3.12}$$

Proposition 3 *If all agents have normal prior distribution at $t = 0$ and "imperfect recall" of past signals, and there is one sophisticated stubborn agent that solves the problem in (3.12), the stubborn s would declare*

$$\theta_s^d = \frac{\mathbf{1}'\boldsymbol{C}\boldsymbol{a}_s}{\boldsymbol{a}_s'\boldsymbol{C}'\boldsymbol{C}\boldsymbol{a}_s + k}\theta_s - \frac{\left(\boldsymbol{C}\left(\boldsymbol{D}(\boldsymbol{I}-\boldsymbol{G})\mathbf{1}\right)\right)'\boldsymbol{a}_s - k}{\boldsymbol{a}_s'\boldsymbol{C}'\boldsymbol{C}\boldsymbol{a}_s + k}\theta^*. \tag{3.13}$$

Moreover, the steady state vector of opinions (beliefs) is

$$\boldsymbol{\mu} = \boldsymbol{C}\left(\left(\boldsymbol{D}(\boldsymbol{I}-\boldsymbol{G}) - \boldsymbol{a}_s\frac{\left(\boldsymbol{C}\left(\boldsymbol{D}(\boldsymbol{I}-\boldsymbol{G})\mathbf{1}\right)\right)'\boldsymbol{a}_s - k}{\boldsymbol{a}_s'\boldsymbol{C}'\boldsymbol{C}\boldsymbol{a}_s + k}\right)\boldsymbol{\theta}^* + \boldsymbol{a}_s\frac{\mathbf{1}'\boldsymbol{C}\boldsymbol{a}_s}{\boldsymbol{a}_s'\boldsymbol{C}'\boldsymbol{C}\boldsymbol{a}_s + k}\theta_s\right) \tag{3.14}$$

Proof. in the Appendix $\square$.

From (3.13), we can study conditions under which the stubborn declare a more extreme opinion than the one that she really has, $|\theta^* - \theta_s^d| > |\theta^* - \theta_s|$.

Without loss of generality we consider only the case where $\theta_s > \theta^* > 0$. From (3.13) we find $\bar{k} = \frac{(\mathbf{1}'\boldsymbol{C}\boldsymbol{a}_s - \boldsymbol{a}_s'\boldsymbol{C}'\boldsymbol{C}\boldsymbol{a}_s)\theta_s - (\boldsymbol{C}(\boldsymbol{D}(\boldsymbol{I}-\boldsymbol{G})\mathbf{1}))'\boldsymbol{a}_s\theta^*}{\theta_s - \theta^*}$ such that:

$$k < \bar{k} \quad \Leftrightarrow \quad \theta_s^d > \theta_s. \tag{3.15}$$

Whenever the cost of lying k overcomes the threshold $\bar{k}$, the sophisticated stubborn induces a lower opinions' distortion than a naive stubborn that declares exactly the opinion she wants to disseminate.

To better understand the implication of Proposition 3, we propose a numerical example.

Example 3: Let us consider the two societies described by Figure 2 (a) and Figure 2 (b). The stubborn in (a) and (b) want to disseminate $\theta_{s_1}^d = 15$ and $\theta_{s_2}^d = 5$, respectively. All agents have "perfect recall" of past signals. The cost of lying for the stubborn is $k = 0.5$. In (a) the stubborn s_1 chooses $\theta_{s_1}^d = 18.6$ and the steady state average opinion is $\boldsymbol{\mu}'\mathbf{1}/n = 11.2900$. On the other hand, in (b) the stubborn s_2 chooses $\theta_{s_2}^d = 5.95$ and the steady state average opinion result to be $\boldsymbol{\mu}'\mathbf{1}/n = 9.4411$. Notice that the same cost of lying $k = 0.5$ is low enough for the more central stubborn s_1 to declare a more extreme opinion ($\theta^d_{s_1} > \theta s_1 > \theta^*$), and high enough for the less central s_2 to declare a less extreme opinion ($\theta^* > \theta^d_{s_2} > \theta s_1$). This suggest us that a more central stubborn has more room to spread misinformation in the society. ■

The main message of this section is that in a "smart society" a naive stubborn – who, given her inability to make a declaration that maximizes her utility, declares her true opinion– is more dangerous than a sophisticated one – who does declare the opinion that maximizes her utility–. In fact, a naive stubborn pursuits is own agenda, to disseminate a certain opinion $\theta_s \neq \theta^*$, no matter on the cost that she faces. On the other hand, a sophisticated stubborn takes into account how costly is to declare a certain opinion. Therefore, the more the society is "smart" the more ineffective is the action of the sophisticated stubborn. We say that a society is more smart than another whenever its the cost of lying k is higher. Therefore, using the term "smart society" we have in mind both a situation in which is the government who can implement policy to enhance cost of lying ($k \uparrow$) or in which is the population's culture that, being is less tolerant to lies, have a higher cost of lying ($k \uparrow$).

These results can provide us with insight about different political campaign strategies and allows us to explain, to a certain extent, the big differences between political statements before and after voting.

We have previously discussed that a policy may have incentives to introduce stubborn in the society to contrast the spread of misinformation and facilitate learning. In Appendix B, we discuss also the competition and the optimal declaration strategy of a stubborn controlled by policy maker to contrast the effect of a sophisticated stubborn in the society.

3.4 Temporary Misinformation

In the previous sections, we have studied social learning only when there are permanent sources of misinformations (stubborn agents), in the society. In this section, we address the study of the speed of convergence to the consensus and the speed of learning of (3.5) in a network without stubborn agents but with temporary misinformation.

In the real world, there exist temporary sources of misinformation (e.g. rumors or fake news) that do not affect the long-run learning but may have important short-run effects. Gratton et al. (2017), for example, show that a bad sender -the spreader of misinformation in our setting-releases information later than good seder, this is mainly due to the fact that in the long-run fake news are unmasked.

Let us consider, for example, a massive diffusion of fake news, or misinformation, in the society that foreruns an election. In such a case, if the learning take place at a too slow pace the temporary distortion can seriously affect the election's outcome. Other examples where the speed of learning play a crucial role are the diffusion of misinformation regarding health, or climate change, issues, the longer the learning process the more serious the damages are. It is important to stress that in this cases the convergence to the consensus is not enough, in fact, agents may agree but still be far from the truth.

With updating rule as in (3.5), if the network is strongly connected (no stubborn), all agents learn the truth, in the long-run (Jadbabaie et al., 2012). However, in the short-run the speed of convergence can play a crucial role. Let us consider the steady state opinions' vector without any stubborn, $\boldsymbol{\mu} = \boldsymbol{\theta}^*$, and we assume that each agent i receive a shock ε_i that represents the diffusion of temporary fake news. The learning process start again by

$$\boldsymbol{\mu}_t = \boldsymbol{\theta}^* + \boldsymbol{\varepsilon} \tag{3.16}$$

We discuss the problem considering the case in which agents are able to "perfect recall" their past signals. Thus, the updating rule is[24]

$$\boldsymbol{\mu}_{t+1} = \boldsymbol{A}\boldsymbol{\mu}_t + \boldsymbol{D}\boldsymbol{\theta}^* \tag{3.17}$$

To study the speed of convergence we avoid to consider the effect of the true state of the world θ^* on agents' updating. Using (3.16), (3.17) becomes

$$\boldsymbol{\mu}_{t+1} = \boldsymbol{A}\boldsymbol{\mu}_t + \boldsymbol{D}(\boldsymbol{\mu}_t + \boldsymbol{\varepsilon}) = \boldsymbol{\mathcal{A}}\boldsymbol{\mu}_t + \boldsymbol{D}\boldsymbol{\varepsilon}$$

$\boldsymbol{\varepsilon}$ represents a small deviation from the steady state. There is consensus when agents' opinion are very similar to each other, regardless of the initial shock. Therefore $\boldsymbol{D}\boldsymbol{\varepsilon}$ is negligible and the consensus time depends only on the convergence of $\boldsymbol{\mathcal{A}}$. The consensus dynamics of (3.17) is equivalent to

$$\boldsymbol{\mu}_{t+1} = \boldsymbol{\mathcal{A}}\boldsymbol{\mu}_t \tag{3.18}$$

Therefore we can use the standard definition of consensus used by Golub and Jackson (2012). Using the standard l^2-norm, we can define consensus time as the time it takes for the distance,

[24] Since we assume a shock at the steady state, under "perfect recall" agents have collect severals signals therefore, they observe directly θ^*.

between average sum of current opinions and steady state opinions of (3.18), to get below ϵ:

Definition [Consensus Time] *The consensus time to $\epsilon > 0$ of a connected graph $\mathcal{G}$ is*

$$CT(\epsilon, \mathcal{G}) := \sup_{\mu_t \in \mathbb{R}^n} \min\{T : ||\boldsymbol{\mathcal{A}}^T \boldsymbol{\mu}_t - \boldsymbol{\mathcal{A}}^\infty \boldsymbol{\mu}_t|| < \epsilon\}$$

Taking the supremum allows us to consider the worst case as a benchmark.

To define the learning time, we still use the l^2-norm. The learning time is the time T, such that the distance between the opinion vector $\boldsymbol{\mu}_{t+T}$ and the true state of the world $\boldsymbol{\theta}^*$, is less then ϵ:

Definition [Learning Time] *The learning time to $\epsilon > 0$ of a connected graph $\mathcal{G}$ is*

$$LT(\epsilon, \mathcal{G}) := \sup_{\mu_t \in \mathbb{R}^n} \min\{T : ||\boldsymbol{\mu}_{t+T} - \boldsymbol{\theta}^*|| < \epsilon\}$$

Again, we use the supremum to consider the worst case as a benchmark.

Notice that in our problem, learning time can be different from consensus time. Agents may have very similar opinions but still be far from the truth and, receiving new pieces of information at each period, they approach $\boldsymbol{\theta}^*$ until they reach it.[25]

We are interested in understanding how the network topology affect the consensus and the learning time. Thus, before to proceed, it is important to define a "bottleneckedness" measure of a network.

Definition [Cheeger Constant] *The Cheeger Constant of the graph $\mathcal{G}(N, \boldsymbol{\mathcal{A}})$ is*

$$\phi(\mathcal{G}) = \min_{S \subseteq N} \sum_{i \in S} \sum_{j \notin S} \frac{a_{ij}}{|S||S^c|}$$

where $S \cup S^c = N$.

The Cheeger constant is a measure of whether or not a graph has a "bottleneck". It quantifies how the network $\mathcal{G}$ can be partitioned in two components. If $\phi(\mathcal{G})$ is small then the network is composed by two sets of vertices with few links between them. On the other hand, if $\phi(\mathcal{G})$ is large, then the network has many links between those two subsets. Moreover, the Cheeger constant is strictly positive if and only if the network is connected.[26]

[25]For example, in the past many people did not believe that smoking cigarette was harmful to health.

[26]The first formulation of Cheeger constant is due to Cheeger (1969), for a deeper discussion of discrete version we remand to Chapter 2 and 6 of Chung (1996).

Defining $1 = \lambda_1^{\mathcal{A}} \geq \lambda_2^{\mathcal{A}} \geq \lambda_n^{\mathcal{A}}$ the eigenvalue of the matrix $\mathcal{A}$.

Proposition 4 *Given the updating rule (3.17) and a network represented by the adjacency matrix $\mathcal{A}$, then for any $\epsilon > 0$ the consensus time $CT(\epsilon, \mathcal{G})$ is in the order of $\lambda_2^{\mathcal{A}}$ exponentially. Moreover,*

$$\frac{\phi(\mathcal{G})^2}{2} \leq 1 - \lambda_2^{\mathcal{A}} \leq 2\phi(\mathcal{G}) \tag{3.19}$$

Proof. in the Appendix. □

This result is standard and consistent with previous literature. In fact, the matrix $\mathcal{A}$ is stochastic and we know from Markov Chain theorem that speed of convergence negatively depends on the magnitude of the second eigenvalue of $\mathcal{A}$. Moreover, if shocks are correlated, $\phi(\mathcal{G})$ is strictly related with the *spectral homophily* measure of Golub and Jackson (2012). The main idea is that the speed of convergence depends on the second largest eigenvalue of $\mathcal{A}$, and according to Cheeger's inequality, the second smallest eigenvalue of the Laplacian matrix, $\lambda_{n-2}^{\mathcal{L}_{\mathcal{A}}}$, is an approximation of the Cheeger constant.[27] We can prove that, in our model, $\lambda_2^{\mathcal{L}_{\mathcal{A}}} = 1 - \lambda_{n-2}^{\mathcal{A}}$, therefore the smaller $\lambda_2^{\mathcal{A}}$ is, the faster the consensus occurs and the more connected the two subsets of nodes are.[28]

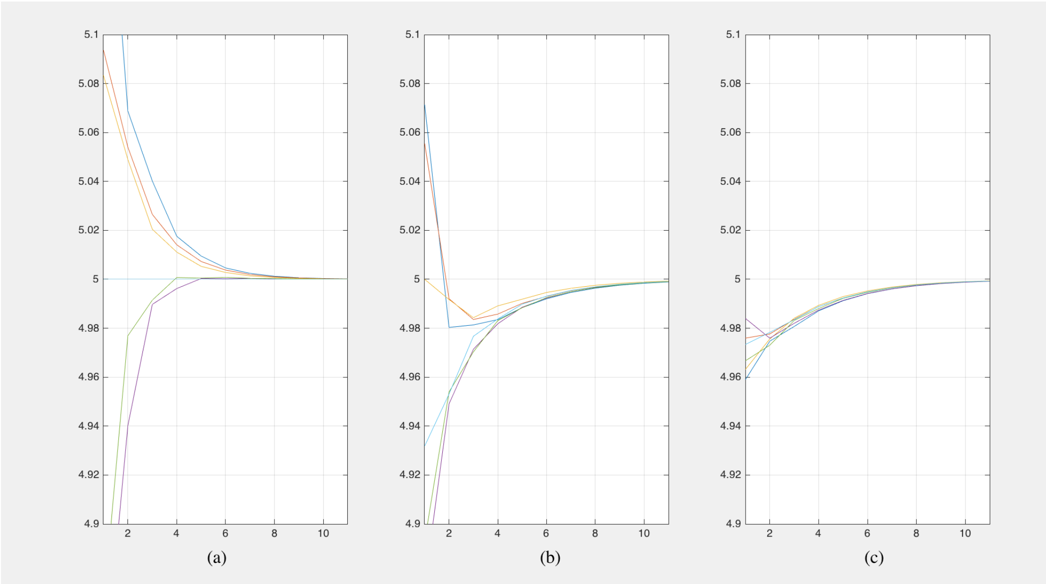
Figure 3 clearly shows how the more the two subsets of nodes are connected, the faster the opinion to converge toward the consensus. In this example, lower "bottleneckedness" means higher homophily, in fact, shocks in $\boldsymbol{\varepsilon}$ are of opposite sign for agents belonging to the two subgroups. We can further see, that a faster consensus does not translate in a faster learning. In Figure 3 we can see that, in this particular example, the society with more "bottleneckedness" (less homophily) is the first to learn the truth in average. Therefore, given our particular updating rule, the speed of learning does not depend only on the "bottleneckedness" (or homophily) of the network as in Golub and Jackson (2012).

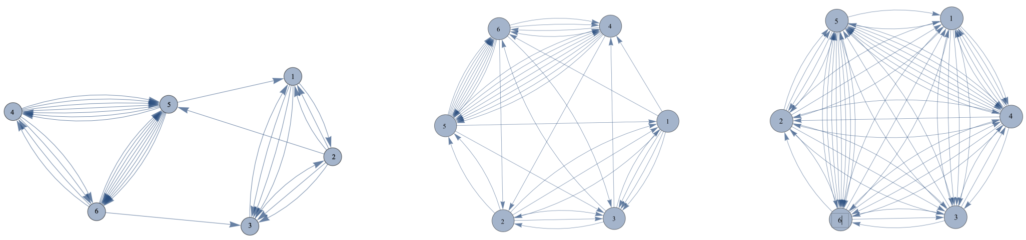
To study the speed of learning the truth, we go back to consider the dynamics in (3.17) where the role of θ^* is explicit. We recall that the steady state of (3.17) is $\boldsymbol{\theta}^*$ and that $\boldsymbol{A}$ is a sub stochastic matrix, in fact $\mathcal{A} = \boldsymbol{A} + \boldsymbol{D}$.

Let us define $\bar{a}_{ii} = \sum_i^{n-1} \frac{a_{ii}}{n-1}$ and ${}_{\min}a_{ii}$ as the average and the minimum among self-weights,

[27] In the appendix we provide technical details and formal definition of Chegeer's inequality and Laplacian matrix. For a deeper discussion about Chegeer's inequality we still remand to Chapter 2 and 6 of Chung (1996).

[28] The second-smallest eigenvalue of the Laplacian matrix $\lambda_2^{\mathcal{L}_{\mathcal{A}}}$ is known as the algebraic connectivity of the graph, and is greater than 0 if and only if the graph described by the adjacency matrix $\mathcal{A}$ is a connected graph.





(a) (b) (c)

Figure 3.3: Convergence of $\boldsymbol{\mu}_t$ to $\theta^* = 5$ in networks with, $\varepsilon_i = 0.25$ for $i = 1, 2, 3$ and $\varepsilon_i = -0.25$ for $i = 4, 5, 6$. (a) $\lambda_2^{\boldsymbol{A}} = 0.9760$, (b) $\lambda_2^{\boldsymbol{A}} = 0.5991$, (c) $\lambda_2^{\boldsymbol{A}} = 0.2283$

respectively. And $\kappa(U)$ as the condition number of the eigenvector basis U.

Proposition 5 *Given the updating rule* (3.17) *and a network represented by the adjacency matrix* $\boldsymbol{\mathcal{A}}$*, then for any* $\epsilon > 0$

$$LT(\epsilon, \mathcal{G}) \leq \lceil \frac{log(\epsilon/(\kappa(U))}{log(|\lambda_1^{\boldsymbol{A}}|)} \rceil \tag{3.20}$$

Moreover,

$$\min_i \{a_{ii}\} \leq 1 - \lambda_1^{\boldsymbol{A}} \leq \bar{a}_{ii} \tag{3.21}$$

Proof. in the Appendix. □

The first part of Proposition 5 shows that, the $LT(\epsilon, \mathcal{G})$ is of the order of the higher eigenvalue of the substochastic matrix $\boldsymbol{A}$ which represent the network without self-weights, and not on the eigenvalue of the full adjacency matrix $\boldsymbol{\mathcal{A}}$. Moreover, positively depends also on the condition number of the eigenvector basis. The second part of Proposition 5 shed light on the role of self-reliance for the speed of learning in our problem. The minimum self-loop $\min_i\{a_{ii}\}$ – which represents the "weakest link" of the learning process, namely the agent that gives less weight to her own private information— is related, through $\lambda_1^{\boldsymbol{\mathcal{A}}}$, to the lower bound of the time of learning. Namely, the higher the minimum individual consideration about the stream of information is, then the faster the learning can be, at its minimum. Thus, it is extremely important how bad informed is the less informed agent, improving the lower level it is possible to ensure a minimum speed of learning. On the other hand, the average self-loop $\bar{a}_{ii}$ – that is the attention that the society gives, on average, to her own private information– can provide information about the upper bound of speed of learning, namely the higher the average individual consideration about the stream of information is, then the faster the learning can be, at its maximum.

Notice that if $a_{ii} = a_{jj} = \alpha$ for all $i, j \in N$ then, from (3.21), $\alpha = 1 - \lambda_1^{\boldsymbol{\mathcal{A}}}$. Moreover from (3.20) $LT(\epsilon, \mathcal{G}) \leq \lceil \frac{log(\epsilon/(\kappa(U))}{log(1-\alpha)} \rceil$.

We have seen, in Proposition 5 and 6, how the "bottleneckedness" (homophily) of the network plays a crucial role in the consensus time, while self-weights are fundamental for what concerns the speed of learning.

We should stress that the speed of learning, even if ultimately depends on $\lambda_1^{\boldsymbol{\mathcal{A}}}$, in the short-run is affected even by other eigenvalues, depending on the magnitude.[29] As an approximation we believe that is enough to consider only $\lambda_1^{\boldsymbol{\mathcal{A}}}, \lambda_2^{\boldsymbol{\mathcal{A}}}$, the two largest eigenvalues.[30] The main intuition is that given two societies with similar levels of self-weights then the different the speed of learning depends on the second largest eigenvalue. Unfortunately, in this case, the interpretation of $\lambda_2^{\boldsymbol{\mathcal{A}}}$ as "bottleneckedness" of the networks is not straightforward.

Corollary 5.1 *Let us consider two symmetric graphs $\mathcal{G}_1(N, \boldsymbol{\mathcal{A}}_1)$ and $\mathcal{G}_2(N, \boldsymbol{\mathcal{A}}_2)$, where $\boldsymbol{\mathcal{A}}_1 = \alpha \boldsymbol{I} + \boldsymbol{A}_1$ and $\boldsymbol{\mathcal{A}}_2 = \alpha \boldsymbol{I} + \boldsymbol{A}_2$, with $\alpha \in [0, 1]$. Given the updating rule (3.17), if $\phi(\mathcal{G}_1) \geq \phi(\boldsymbol{G}_2)$ then for any $\epsilon > 0$*

- $CT(\epsilon, \mathcal{G}_1) \leq CT(\epsilon, \mathcal{G}_2)$
- $LT(\epsilon, \mathcal{G}_1) \leq LT(\epsilon, \mathcal{G}_2)$

Proof. in the Appendix □

[29] We remand to the proof of Proposition 6 for details.

[30] Notice, in fact, that $\boldsymbol{\mathcal{A}}$ is row sub-stochastic therefore all the eigenvalues are less then 1 and most of them vanish after very few iterations.

This corollary discusses the particular case where all agents in two different symmetric networks give the same weight to the information that they receive. It shows that the ordering of second largest eigenvalues of adjacency matrices relative to the network without self-loops $\lambda_2^{\mathcal{A}_1}$ and $\lambda_2^{\mathcal{A}_2}$ is the same the ordering of "bottleneckedness" measures $\phi(\mathcal{G}_1)$ and $\phi(\mathcal{G}_2)$. In Figure 4, we can observe a numerical example where agents give the same weights to their private information, in such a case, a higher homophily leads to a faster consensus and learning.

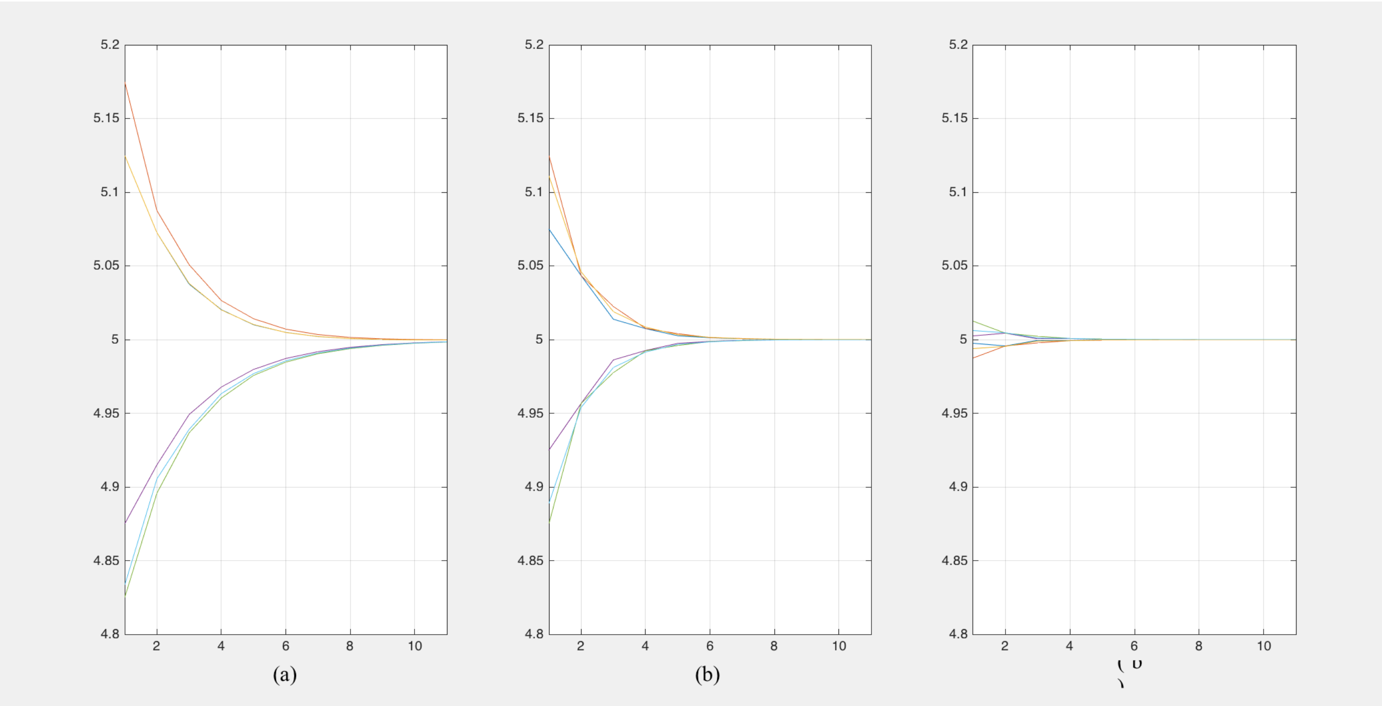


Figure 3.4: Convergence of $\boldsymbol{\mu}_t$ to $\theta^* = 5$ in networks with equal self-loops $\alpha = 0.3$, $\theta^* = 5$, $\varepsilon_i = 0.25$ for $i = 1, 2, 3$ and $\varepsilon_i = -0.25$ for $i = 4, 5, 6$. (a) $\lambda_2^{\boldsymbol{A}} = 0.5846$, (b) $\lambda_2^{\boldsymbol{A}} = 0.4209$, (c) $\lambda_2^{\boldsymbol{A}} = -0.1$

From Proposition 4 and 5 and Corollary 5.1 we deduce that, if agents receive a continuous stream of new information about the true state of the world then, the speed of learning and the speed of convergence to the consensus are, in general, different. The first one mainly depends on the self-reliance of agents, namely how much they care about the information that they receive. The second strongly depends on the "bottleneckedness" of the network and therefore, when shocks are correlated with network structure, on the homophily. If all agents have the same fixed reliance on private signals and the network is symmetric, then a smaller "bottleneckedness" (higher homophily) leads both to a lower consensus and learning time.

Comparing results with Golub and Jackson (2012) the main insight is that, if agents receive new information at each time then the speed of learning is not necessarily directly correlated to homophily. For example, an increase in the number of connections among agents belonging to two different subset of the network[31] may translate into a lower speed of learning (as in Figure 3), due to the rescaling of self-weights.[32] This strongly relies on the fact that agents receive constant flows of information, unlike in a DeGroot like social learning framework as Golub and Jackson (2012).

From a policymaker point of view, these results suggest an increase in the network density does not directly translate into a faster learning of the truth. Moreover, policies that aim to increase social interaction and decreasing the homophily to facilitate agents' learning, may succeed only if agents belonging to different groups have the same level "good information". On the contrary, the learning of better informed agents is slowed down by others and this reduce the speed of learning of the whole society.

3.5 Conclusion

This paper addresses the problem of the spread of misinformation in a social network where agents interact to learn an underlying state of the world with a non-Bayesian social learning process. The main difference with the standard naive social learning, as in Golub and Jackson (2010), is the continuous stream of new signals that agents receive at each period. This implies a stronger connection to the truth in the learning process. Considering the permanent misinformation or opinion manipulation pursuit by "prominent" agents we show that despite receiving new signals every period, agents are not able to learn the underlying state of the word nor to reach a consensus. This depends on the fact that the network is not strongly connected due to the presence of "prominent agents" who behave as stubborn. Differently from the benchmark of DeGroot model, if agents receive signals at each period, the steady state agents' opinion does not depends only on the stubborn opinion and the centrality of agents connected to stubborn but also on the true state of the world and on agents' self-reliance. We, therefore, introduce a novel centrality measure the "updating centrality" that, in the case of "perfect recall" of past signals, corresponds to the Katz-Bonachic centrality. We further characterize the optimal action of a stubborn who want to manipulate the opinion dynamics showing his relationship with the cost of disseminating misinformation. At the end, we discuss the consensus and learning time after an exogenous shock that temporarily moves opinions away from the steady state. We prove that the speed of reaching the consensus inversely depends on the "bottleneckedness", and thus the homophily, of the underlying network, while the speed of convergence toward the

[31] That, as said, correspond to a decrease in the level of homophily if shocks are correlated.

[32] Given that $\sum_j a_{ij} = 1$ for all $j \in I$, if for i a_{ik} increase, where $k \in I/\{i\}$, then the sum of others a_{ij}, for $j \in I/\{k\}$ should decrease of the same amount, thus a_{ii} can decrease too.

truth mainly depends on the strength of self-weights.

A potential interesting extension for the future is to study a similar framework in a society where agents have the possibility to choose their connection. The probable result is that if stubborn agents are not very central and agents strongly rely on their private signals, then stubborn agents end up to be isolated in the long-run and agents reach a full learning. On the other hand, if stubborn is central and agents have low reliance on private signals can arise different isolated communities where agents as opinion very close to the stubborn. In this framework, it is interesting to investigate if a sophisticated stubborn would declare more extreme opinions when she is central or peripheral.

Appendix A

Proofs of Propositions

Lemma 1

Before presenting the proof of Proposition 1 and Corollary 1.1 we have to state and proof the following lemma.[33]

Lemma 1 (Jadbabaie et al., 2012) *Let $\boldsymbol{\mathcal{A}}_s$ denote the matrix of social interactions. The sequence $\sum_{i=1}^{n} \nu_i p_{i,t}(\theta^*)$ converges $\mathbb{P}^*$ - almost surely as $t \to \infty$, where $\boldsymbol{\nu}$ is any non-negative left eigenvector of $\boldsymbol{\mathcal{A}}_s$ corresponding to its unit eigenvalue.*

Proof.

Notice that since $\boldsymbol{\mathcal{A}}_s$ is stochastic its largest eigenvalue is $\lambda_1 = 1$. Moreover, always exists $\boldsymbol{\nu}$ a non-negative left eigenvector corresponding to the eigenvalue $\lambda_1 = 1$. We define $\boldsymbol{p}_{t+1}(\theta)$ the vector of probabilistic belief of all $i \in N$ and $p_{s1,t+1}(\theta), p_{s2,t+1}(\theta)$ the stubborn beliefs. Since stubborn do not revise their beliefs we can consider only the probabilistic belief's updating rule (1) for all $i \in N$ (evaluated at θ^*)

[33] We adapt the proof from Jadbabaie et al. (2012), the only difference is the presence of stubborn agents that makes $\boldsymbol{\mathcal{A}}_s$ not to be strongly-connected.

$$\boldsymbol{p}_{t+1}(\theta^*) = \boldsymbol{\mathcal{A}}\boldsymbol{p}_t(\theta^*)+\sum_{i=1}^{n} p_{i,t}(\theta^*)a_{ii}\left(\frac{l(\omega_{i,t+1}|\theta^*)}{\int_\Theta l(\omega_{i,t+1}\theta^*)p_{i,t}(\theta^*)d\theta}-1\right)+\sum_{i=1}^{n}\left[a_{i,s_1}p_{s_1}(\theta^*)+a_{i,s_2}p_{s_2}(\theta^*)\right] \tag{3.22}$$

$\boldsymbol{\mathcal{A}}$ is a stochastic matrix, therefore the largest eigenvalue is $\lambda_1 = 1$, we denote with $\boldsymbol{\nu}$ the eigenvector corresponding to λ_1. Notice that all element in $\boldsymbol{\nu}$ are non negative and $\boldsymbol{\nu}'\boldsymbol{\mathcal{A}} = \boldsymbol{\nu}'\lambda_1$. Moreover, for a generic stubborn $p_s(\theta) = 0$ for all $\theta \neq \theta_s$.

Let us multiply both sides of (20) by $\boldsymbol{\nu}'$,

$$\boldsymbol{\nu}'\boldsymbol{p}_{t+1}(\theta^*) = \boldsymbol{\nu}'\boldsymbol{\mathcal{A}}\boldsymbol{p}_t(\theta^*)+\sum_{i=1}^{n} \nu_i p_{i,t}(\theta^*)a_{ii}\left(\frac{l(\omega_{i,t+1}|\theta^*)}{\int_\Theta l(\omega_{i,t+1}\theta^*)p_{i,t}(\theta^*)d\theta}-1\right)+\underbrace{\sum_{i=1}^{n}\nu_i\left[a_{i,s_1}p_{s_1}(\theta^*)+a_{i,s_2}p_{s_2}(\theta^*)\right]}_{=0}$$

then we take the expectation $\mathbb{E}$ associated with measure $\mathbb{P}^*$ with respect the filtration $\mathcal{F}_t$

$$\mathbb{E}\left[\boldsymbol{\nu}'\boldsymbol{p}_{t+1}(\theta^*)|\mathcal{F}_t\right] = \boldsymbol{\nu}'\boldsymbol{p}_t(\theta^*) + \sum_{i=1}^{n}\nu_i p_{i,t}(\theta^*)a_{ii}\mathbb{E}\left[\left(\frac{l(\omega_{i,t+1}|\theta^*)}{\int_\Theta l(\omega_{i,t+1}\theta^*)p_{i,t}(\theta^*)d\theta}-1\right)|\mathcal{F}_t\right]$$

Jensen's inequality implies that

$$\mathbb{E}\left[\left(\frac{l(\omega_{i,t+1}|\theta^*)}{\int_\Theta l(\omega_{i,t+1}\theta^*)p_{i,t}(\theta^*)d\theta}\right)|\mathcal{F}_t\right] \geq \left(\mathbb{E}\left[\left(\frac{l(\omega_{i,t+1}|\theta^*)}{\int_\Theta l(\omega_{i,t+1}\theta^*)p_{i,t}(\theta^*)d\theta}\right)^{-1}|\mathcal{F}_t\right]\right)^{-1} = 1$$

then

$$\mathbb{E}\left[\boldsymbol{\nu}'\boldsymbol{p}_{t+1}(\theta^*)|\mathcal{F}_t\right] = \boldsymbol{\nu}'\boldsymbol{p}_t(\theta^*) + \sum_{i=1}^{n}\nu_i p_{i,t}(\theta^*)a_{ii}\underbrace{\mathbb{E}\left[\left(\frac{l(\omega_{i,t+1}|\theta^*)}{\int_\Theta l(\omega_{i,t+1}\theta^*)p_{i,t}(\theta^*)d\theta}-1\right)|\mathcal{F}_t\right]}_{\geq 0}$$

Therefore

$$\mathbb{E}\left[\boldsymbol{\nu}'\boldsymbol{p}_{t+1}(\theta^*)|\mathcal{F}_t\right] \geq \boldsymbol{\nu}'\boldsymbol{p}_t(\theta^*)$$

Thus, since $\boldsymbol{\nu}'\boldsymbol{p}_t(\theta^*)$ is a submartingale with respect $\mathcal{F}_t$ then it converges $\mathbb{P}^*$ - almost surely.

□

Proof of Proposition 1

If the prior distribution of agents at $t = 0$ is normal then it remain normal for each t, in fact the bayesian posterior of a normal remains normal and a sum of two normal distributed random variable remains normally distributed. Moreover the mean of the bayesian posterior distribution is

$$\bar{\beta}_{i,t+1} = \int_{\Theta} \theta \beta_{i,t+1} d\theta = \gamma_{i,t}\mu_{i,t} + (1 - \gamma_{i,t})\omega_{i,t+1} \tag{3.23}$$

where $\omega_{i,t+1}$ is the signal received by agent i at $t + 1$. Thus (4) becomes

$$\boldsymbol{\mu}_{t+1} = \boldsymbol{D}(\boldsymbol{G}_t\boldsymbol{\mu}_t + (\boldsymbol{I} - \boldsymbol{G}_t)\boldsymbol{\omega}_t) + \boldsymbol{A}\boldsymbol{\mu}_t + \boldsymbol{a}_{s1}\theta_{s_1} + \boldsymbol{a}_{s2}\theta_{s_2} \tag{3.24}$$

From Lemma 1, we know that the probability distribution converges, $p_{i,t}(\theta) \to p_i(\theta)$ therefore both the mean and the precision converges almost surely $\mu_{i,t} \to \mu_i$, $\tau^p_{i,t} \to \tau^p_i$ and thus $\gamma_{i,t} \to \gamma_i$. Moreover, we recall that $E[\boldsymbol{\omega}] = \boldsymbol{\theta}^*$. Thus, by the law of large numbers at the steady state

$$\boldsymbol{\mu} = \boldsymbol{D}(\boldsymbol{G}\boldsymbol{\mu} + (\boldsymbol{I} - \boldsymbol{G})\boldsymbol{\theta}^*) + \boldsymbol{A}\boldsymbol{\mu} + \boldsymbol{a}_{s1}\theta_{s_1} + \boldsymbol{a}_{s2}\theta_{s_2}$$

That lead us to

$$\Rightarrow \boldsymbol{\mu} = \underbrace{(\boldsymbol{I} - \boldsymbol{D}\boldsymbol{G} - \boldsymbol{A})^{-1}}_{\boldsymbol{C}}(\boldsymbol{D}(\boldsymbol{I} - \boldsymbol{G})\boldsymbol{\theta}^* + \boldsymbol{a}_{s1}\theta_{s_1} + \boldsymbol{a}_{s2}\theta_{s_2})$$

□

Proof of Corollary 1.1

If agents are able to recall all the past signals after T period they will compute their bayesian posterior using all the T signals, therefore

$$\gamma_{i,T} = \frac{\tau^p_{i,T}}{\tau^p_{i,T} + T * \tau^\omega}$$

since

$$\lim_{T\to\infty} \gamma_{i,T} = 0$$

Then all elements of matrix $\boldsymbol{G}$ approach zero time to time. At the steady state

$$\Rightarrow \boldsymbol{\mu} = (\boldsymbol{I} - \boldsymbol{A})^{-1}(\boldsymbol{D}\boldsymbol{\theta}^* + \boldsymbol{a}_{s1}\theta_{s_1} + \boldsymbol{a}_{s2}\theta_{s_2})$$

□

Proof of Proposition 2

In order to prove Proposition 2 we state this well known linear algebra result.

Sherman-Morrison Formula. (Sherman and Morrison, 1950) *Let $\boldsymbol{B}$ be a nonsingular n-dimensional real matrix, and $\boldsymbol{u}, \boldsymbol{v}$ two real n-dimensional column vectors such that $1+\boldsymbol{v}'A^{-1}\boldsymbol{u} \neq 0$. Then,*

$$(\boldsymbol{B}+\mathbf{u}\mathbf{v}')^{-1} = \boldsymbol{B}^{-1} - \frac{\boldsymbol{B}^{-1}\mathbf{u}\mathbf{v}'\boldsymbol{B}^{-1}}{1+\mathbf{v}'\boldsymbol{B}^{-1}\mathbf{u}}$$

Since $\boldsymbol{G}$ is a diagonal matrix and α_{ii} is assumed to be zero, then we can apply the Sherman-Morrison Formula, where $\boldsymbol{B} = (\boldsymbol{I} - \boldsymbol{DG} - \boldsymbol{A})$, $\mathbf{u} = \mathbf{e}_i$ and $\mathbf{v} = \mathbf{g}_i$.

We study the general case. Let us consider the steady state opinion vector.

$$\boldsymbol{\mu} = (\boldsymbol{I} - \boldsymbol{DG} - \boldsymbol{A})^{-1}\left(\boldsymbol{D}(\boldsymbol{I} - \boldsymbol{G})\boldsymbol{\theta}^* + \boldsymbol{a}_{s1}\theta_{s1} + \boldsymbol{a}_{s2}\theta_{s2}\right)$$

If agents i increase the interaction with the stubborn s_1 of α then will decrease proportionally the interaction of other agents as captured by the vector $\boldsymbol{g}_i$. The new interaction matrix is $\hat{\boldsymbol{\mathcal{A}}} = \boldsymbol{\mathcal{A}} - \mathbf{e}_i\mathbf{g}_i'$. Since $\mathbf{e}_i(\mathbf{g}_i)'$ does not affect the main diagonal $\boldsymbol{D}$, the new opinion vector. $\hat{\boldsymbol{\mu}}$ is

$$\Rightarrow \hat{\boldsymbol{\mu}} = (\boldsymbol{I} - \boldsymbol{DG} - \boldsymbol{A} + \mathbf{e}_i\mathbf{g}_i')^{-1}\left(\boldsymbol{D}(\boldsymbol{I} - \boldsymbol{G})\boldsymbol{\theta}^* + (\boldsymbol{a}_{s1} + \alpha\mathbf{e}_i)\theta_{s1} + \boldsymbol{a}_{s2}\theta_{s2}\right)$$

By Sherman-Morrison Formula we know that

$$(\boldsymbol{I} - \boldsymbol{DG} - \boldsymbol{A} + \mathbf{e}_i\mathbf{g}_i')^{-1} = (\boldsymbol{I} - \boldsymbol{DG} - \boldsymbol{A})^{-1} - \frac{(\boldsymbol{I} - \boldsymbol{DG} - \boldsymbol{A})^{-1}\mathbf{e}_i\mathbf{g}_i'(\boldsymbol{I} - \boldsymbol{DG} - \boldsymbol{A})^{-1}}{1 + \mathbf{g}_i'(\boldsymbol{I} - \boldsymbol{DG} - \boldsymbol{A})^{-1}\mathbf{e}_i}$$

The first term on the right side is $\boldsymbol{C}$. We name $\boldsymbol{X}$ the second one. Thus we obtain

$$\Rightarrow \hat{\boldsymbol{\mu}} = (\boldsymbol{C} - \boldsymbol{X})\left(\boldsymbol{D}(\boldsymbol{I} - \boldsymbol{G})\boldsymbol{\theta}^* + (\boldsymbol{a}_{s1} + \alpha\mathbf{e}_i)\theta_{s1} + \boldsymbol{a}_{s2}\theta_{s2}\right)$$

$$\Rightarrow \hat{\boldsymbol{\mu}} = \underbrace{\boldsymbol{C}\left(\boldsymbol{D}(\boldsymbol{I} - \boldsymbol{G})\boldsymbol{\theta}^* + \boldsymbol{a}_{s1}\theta_{s1} + \boldsymbol{a}_{s2}\theta_{s2}\right)}_{\boldsymbol{\mu}} + \underbrace{\left((\boldsymbol{C} - \boldsymbol{X})\alpha\mathbf{e}_i - \boldsymbol{X}\boldsymbol{a}_{s1})\theta_{s1} - \boldsymbol{X}(\boldsymbol{D}(\boldsymbol{I} - \boldsymbol{G})\boldsymbol{\theta}^* + \boldsymbol{a}_{s2}\theta_{s2})\right)}_{\Delta\boldsymbol{\mu}}$$

If we want to see the effect of creating the first link of intensity α with a stubborn for a generic

i it is enough to consider a interaction matrix where $\boldsymbol{a}_{s1} = \boldsymbol{a}_{s2} = 0$, thus

$$\Rightarrow \hat{\boldsymbol{\mu}} = \underbrace{\boldsymbol{C}\left(\boldsymbol{D}(\boldsymbol{I}-\boldsymbol{G})\boldsymbol{\theta}^*\right)}_{\boldsymbol{\mu}} + \underbrace{(\boldsymbol{C}-\boldsymbol{X})\alpha\mathbf{e}_i\theta_{s1} - \boldsymbol{X}\boldsymbol{D}(\boldsymbol{I}-\boldsymbol{G})\boldsymbol{\theta}^*)}_{\Delta\boldsymbol{\mu}}$$

$$\Rightarrow \hat{\boldsymbol{\mu}} = (\boldsymbol{C}-\boldsymbol{X})(\alpha\mathbf{e}_i\theta_{s1} + \boldsymbol{D}(\boldsymbol{I}-\boldsymbol{G})\boldsymbol{\theta}^*)$$

We can see that

$$X = \frac{(\boldsymbol{I}-\boldsymbol{D}\boldsymbol{G}-\boldsymbol{A})^{-1}\mathbf{e}_i\mathbf{g}_i'(\boldsymbol{I}-\boldsymbol{D}\boldsymbol{G}-\boldsymbol{A})^{-1}}{1+\mathbf{g}_i'(\boldsymbol{I}-\boldsymbol{D}\boldsymbol{G}-\boldsymbol{A})^{-1}\mathbf{e}_i}$$

measures the variation of the updating centrality after introducing a stubborn in the society.

□

Proof of Proposition 3

Let us consider the maximization problem described in (3.12)

$$\max_{\theta_s^d} \; u_s(\boldsymbol{\mu}): \; -\left(\boldsymbol{\mu}-\mathbf{1}\theta_s\right)^2 - k(\theta^*-\theta_s^d)^2$$

Expanding the first term we obtain

$$\max_{\theta_s^d} \; -\boldsymbol{\mu}'\boldsymbol{\mu} + 2(\mathbf{1}\theta_s)'\boldsymbol{\mu} - (\mathbf{1}\theta_s)'(\mathbf{1}\theta_s) - k(\theta^*-\theta_s^d)^2$$

We can see that this problem is equivalent to

$$\max_{\theta_s^d} \; -\boldsymbol{\mu}'\boldsymbol{\mu} + 2(\mathbf{1}\theta_s)'\boldsymbol{\mu} - k(\theta^*-\theta_s^d)^2 \tag{3.25}$$

With only one sophisticated stubborn agent that declares θ_s^d the steady state opinion dynamics is

$$\boldsymbol{\mu} = \boldsymbol{C}\left(\boldsymbol{D}(\boldsymbol{I}-\boldsymbol{G})\boldsymbol{\theta}^* + \boldsymbol{a}_s\theta_s^d\right) \tag{3.26}$$

Substituting (3.26) in (3.25) we obtain

$$\max_{\theta_s^d} \; -2\left(\boldsymbol{C}\left(\boldsymbol{D}(\boldsymbol{I}-\boldsymbol{G})\boldsymbol{\theta}^*\right)\right)'\boldsymbol{a}_s\theta_s^d - \boldsymbol{a}_s'\boldsymbol{C}'\boldsymbol{C}\boldsymbol{a}_s\theta_s^{d2} + 2(\mathbf{1}\theta_s)'\boldsymbol{C}\boldsymbol{a}_s\theta_s^d - k(\theta^*-\theta_s^d)^2$$

We solve the First Order Condition:

$$-2\left(\boldsymbol{C}\left(\boldsymbol{D}(\boldsymbol{I}-\boldsymbol{G})\boldsymbol{\theta}^*\right)\right)'\boldsymbol{a}_s - 2\boldsymbol{a}_s'\boldsymbol{C}'\boldsymbol{C}\boldsymbol{a}_s\theta_s^d + 2(\mathbf{1}\theta_s)'\boldsymbol{C}\boldsymbol{a}_s + 2k(\theta^*-\theta_s^d) = 0$$

$$\Rightarrow \theta_s^d = \frac{-\left(\boldsymbol{C}\left(\boldsymbol{D}(\boldsymbol{I}-\boldsymbol{G})\boldsymbol{\theta}^*\right)\right)'\boldsymbol{a}_s + (\boldsymbol{1}\theta_s)'\boldsymbol{C}\boldsymbol{a}_s + k\theta^*}{\boldsymbol{a}_s'\boldsymbol{C}'\boldsymbol{C}\boldsymbol{a}_s + k}$$

Therefore we get the optimal declaration for an optimizing stubborn.

$$\theta_s^d = \frac{\boldsymbol{1}'\boldsymbol{C}\boldsymbol{a}_s}{\boldsymbol{a}_s'\boldsymbol{C}'\boldsymbol{C}\boldsymbol{a}_s + k}\theta_s - \frac{\left(\boldsymbol{C}\left(\boldsymbol{D}(\boldsymbol{I}-\boldsymbol{G})\boldsymbol{1}\right)\right)'\boldsymbol{a}_s - k}{\boldsymbol{a}_s'\boldsymbol{C}'\boldsymbol{C}\boldsymbol{a}_s + k}\theta^*$$

Substituting θ_s^d in the steady state opinion vector equation we get exactly (3.14).

□

Proof of Proposition 4

To prove that the consensus time of (3.17) is in the order of $\lambda_2^{\boldsymbol{\mathcal{A}}}$ exponentially we use the following well-know theorem

Theorem 6 (Perron-Frobenius) *Let the eigenvectors be chosen so that* $\boldsymbol{\nu}_i'\boldsymbol{v}_i = 1$*, where* $\boldsymbol{\nu}_i'$ *is the left eigenvector and* $\boldsymbol{v}_i$ *is the right eigenvector.* λ_1 *and* λ_2 *the first and second largest eigenvalue and* r_2 *the algebraic multiplicity associated with* λ_2*. Then we get*

$$\boldsymbol{\mathcal{A}}^n = (\lambda_1^{\boldsymbol{\mathcal{A}}})^n \boldsymbol{v}_i\boldsymbol{\nu}_i' + o(n^{r_2-1}|\lambda_2^{\boldsymbol{\mathcal{A}}}|^n)$$

Corollary 6.1 *Since* $\boldsymbol{\mathcal{A}}$ *is a stochastic aperiodic matrix* $\lambda_1^{\boldsymbol{\mathcal{A}}} = 1$ *and* $\boldsymbol{v} = \boldsymbol{1}$ *if the algebraic multiplicity associated with* $\lambda_2^{\boldsymbol{\mathcal{A}}}$ r_2 *is equal to 1, then*

$$\boldsymbol{\mathcal{A}}^n = \boldsymbol{1}\boldsymbol{\nu}_i' + o(|\lambda_2^{\boldsymbol{\mathcal{A}}}|^n)$$

a smaller second-largest eigenvalue directly corresponds to a higher rate of convergence. $\Rightarrow$ $CT(\epsilon, \mathcal{G})$ is in the order of $\lambda_2^{\boldsymbol{\mathcal{A}}}$ exponentially.

Before to prove inequality (3.19) we have to introduce the definition of Laplacian matrix and the result known as Cheeger's inequality.[34]

Definition (Laplacian Matrix) *A matrix* $\boldsymbol{\mathcal{L}} = (l_{ij}) \in \mathbb{R}^{n\times n}$ *is a Laplacian Matrix* $\boldsymbol{\mathcal{L}}$ *iff*

1. $l_{i,j} \leq 0, \quad j \neq i$
2. $\sum_{j=1}^{n} l_{i,j} = 0, \quad i = 1, 2, ..., n$

[34] We refer to Agaev and Chebotarev (2005) for the discussion about non-symmetric Laplacian matrices and to Chung (1996) for Cheeger's inequality.

The Laplacian Matrix can be computed as the difference between the diagonal degree matrix and the adjacency matrix.

Cheeger's Inequality (Chung, 1996) *If $\lambda_2^{\mathcal{L}}$ is the second smallest eigenvalue of the Laplacian of the graph $\mathcal{G}(N, \boldsymbol{\mathcal{A}})$, then:*

$$\frac{\phi(\mathcal{G})^2}{2} \leq \lambda_{n-2}^{\mathcal{L}^{\boldsymbol{\mathcal{A}}}} \leq 2\phi(\mathcal{G})$$

$\boldsymbol{\mathcal{A}}$ is a row-stochastic matrix, thus his Laplacian is $\mathcal{L}^{\boldsymbol{\mathcal{A}}} = \boldsymbol{I} - \boldsymbol{\mathcal{A}}$ and its the second smallest eigenvalue is nothing but $\lambda_{n-2}^{\mathcal{L}^{\boldsymbol{\mathcal{A}}}} = 1 - \lambda_2^{\boldsymbol{\mathcal{A}}}$, where $\lambda_2^{\boldsymbol{\mathcal{A}}}$ is second largest eigenvalue of the adjacency matrix. Therefore

$$\Rightarrow \frac{\phi(\mathcal{G})^2}{2} \leq 1 - \lambda_2^{\boldsymbol{\mathcal{A}}} \leq 2\phi(\mathcal{G})$$

□

Proof of Proposition 5

Iterating the process (3.17) we get

$$\boldsymbol{\mu}_{t+T} = \boldsymbol{A}^T \boldsymbol{\mu}_t + \sum_{i=0}^{T-1} \boldsymbol{A}^i \boldsymbol{D} \boldsymbol{\theta}^*$$

$$\boldsymbol{\mu}_{t+T} = \boldsymbol{A}^T \boldsymbol{\mu}_t + \frac{\boldsymbol{I} - \boldsymbol{A}^T}{\boldsymbol{I} - \boldsymbol{A}} \boldsymbol{D} \boldsymbol{\theta}^*$$

$$\boldsymbol{\mu}_{t+T} = \boldsymbol{A}^T \boldsymbol{\mu}_t + (\boldsymbol{I} - \boldsymbol{A}^T) \boldsymbol{\theta}^*$$

$$\boldsymbol{\mu}_{t+T} - \boldsymbol{\theta}^* = \boldsymbol{A}^T (\boldsymbol{\mu}_t - \boldsymbol{\theta}^*)$$

Thus depends on how fast is $\boldsymbol{A}^T \to 0$. If $\boldsymbol{A}$ is diagonalizable then, we can define $\boldsymbol{U}$ as the eigenvector matrix and $\boldsymbol{\Lambda}$ as the diagonal matrix with eigenvalues on its main diagonal.

$$\boldsymbol{A} = \boldsymbol{U} \boldsymbol{\Lambda} \boldsymbol{U}^{-1}$$

is the eigendecomposition of $\boldsymbol{A}$. Therefore iterating it T times we get

$$\boldsymbol{A}^T = \boldsymbol{U} \boldsymbol{\Lambda}^T \boldsymbol{U}^{-1}$$

$$
\begin{aligned}
\|\boldsymbol{\mu}_{t+T} - \boldsymbol{\theta}^{*}\| &= \|\boldsymbol{A}^{T}(\boldsymbol{\mu}_{t} - \boldsymbol{\theta}^{*})\| \\
&= \|\boldsymbol{U}\boldsymbol{\Lambda}^{T}\boldsymbol{U}^{-1}(\boldsymbol{\mu}_{t} - \boldsymbol{\theta}^{*})\| \\
&\leq \sum_{j=1}^{n} |\lambda_{j}^{\boldsymbol{A}}|^{T} \|\boldsymbol{U}\| \|\boldsymbol{U}^{-1}\| \|(\boldsymbol{\mu}_{t} - \boldsymbol{\theta}^{*})\| \\
&\leq |\lambda_{1}^{\boldsymbol{A}}|^{T} \|\boldsymbol{U}\| \|\boldsymbol{U}^{-1}\|
\end{aligned}
$$

Where the last inequality stem from (3.16). Moreover, if

$$
T \geq \frac{log(\epsilon/(\kappa(U)))}{log(|\lambda_{1}^{\boldsymbol{A}}|)}
$$

Then

$$
\|\boldsymbol{\mu}_{t+T} - \boldsymbol{\theta}^{*}\| \leq \epsilon
$$

Therefore

$$
LT(\epsilon, \mathcal{G}) \leq \lceil \frac{log(\epsilon/(\kappa(U))}{log(|\lambda_{1}^{\boldsymbol{A}}|)} \rceil
$$

The $LT(\epsilon, \mathcal{G})$ depends on eigenvalues and eigenvectors. We can, thus, conclude that $LT(\epsilon, \mathcal{G})$ is of the order of the higher eigenvalue, exponentially. And positively depends on the condition number of the eigenvector basis. Notice that all eigenvalues have a short-run effect, that decays over time according to the their absolute values.

We now prove the inequality (3.21) in Proposition 7. Let us define as $\bar{d}^{\boldsymbol{A}}$ and $d_{\max}^{\boldsymbol{A}}$ as the average and the maximum degree, respectively

- Lower bound

  Using the Rayleigh quotient (Horn and Johnson, 1985)

$$
\frac{\mathbf{1}'\boldsymbol{A}\mathbf{1}}{\mathbf{1}'\mathbf{1}} = \frac{\sum_{ij} a_{ij}}{n} = \frac{\sum_{i} d_{i}^{\boldsymbol{A}}}{n} = \bar{d}^{\boldsymbol{A}} \leq \lambda_{1}^{\boldsymbol{A}} \tag{3.27}
$$

- Upper bound

  Let $\boldsymbol{\nu}_{1}$ be an eigenvector belonging to $\lambda_{1}^{\boldsymbol{A}}$ and ν_{1i} be the entry with largest absolute value. Then

$$
\lambda_{1}^{\boldsymbol{A}}|\nu_{1i}| = \sum_{j} a_{ij}|\nu_{1i}| \leq d_{\max}^{\boldsymbol{A}}|\nu_{1i}|
$$

$$\Rightarrow \lambda_1^{\boldsymbol{A}} = \sum_j a_{ij} \leq d_{\max}^{\boldsymbol{A}} \tag{3.28}$$

Putting together (3.27) and (3.28) we finally get

$$\Rightarrow \bar{d}^{\boldsymbol{A}} \leq \lambda_1^{\boldsymbol{A}} \leq d_{\max}^{\boldsymbol{A}}$$

Since $\boldsymbol{A} = \boldsymbol{\mathcal{A}} - \boldsymbol{D}$, then the degree of a generic agent i is $d_i = \sum_i a_{ij} = 1 - a_{ii}$. Thus, the average degree $\bar{d}^{\boldsymbol{A}} = 1 - \sum_i \frac{a_{ii}}{n} = 1 - \bar{a}_{ii}$ and the maximum degree of $\boldsymbol{A}$ is $d_{\max}^{\boldsymbol{A}} = \max_i\{1 - a_{ii}\} = 1 - \min_i\{a_{ii}\}$. Therefore, for a graph described by the adjacency matrix $\boldsymbol{A}$

$$1 - \bar{a}_{ii} \leq \lambda_1^{\boldsymbol{A}} \leq 1 - \min_i\{a_{ii}\}$$

□

Proof of Corollary 5.1

In general, by Cheeger's inequality we have that

$$2\mathcal{L}_{n-2}^{\boldsymbol{\mathcal{A}}} \leq \phi(\mathcal{G}) \leq 2\sqrt{\lambda_{n-2}^{\mathcal{L}^{\boldsymbol{\mathcal{A}}}}}$$

Namely, $\phi(\mathcal{G})$ is increasing in $\lambda_{n-2}^{\mathcal{L}^{\boldsymbol{\mathcal{A}}}} = 1 - \lambda_2^{\boldsymbol{\mathcal{A}}}$. Therefore we can conclude that

$$\phi(\mathcal{G}_1) \geq \phi(\mathcal{G}_2) \Longrightarrow \lambda_1^{\boldsymbol{\mathcal{A}}} \leq \lambda_2^{\boldsymbol{\mathcal{A}}} \Longrightarrow CT(\epsilon, \mathcal{G}_1) \leq CT(\epsilon, \mathcal{G}_2) \tag{3.29}$$

Where the last implication stem from Proposition 6.

If $\boldsymbol{\mathcal{A}}_1 = \alpha \boldsymbol{I} + \boldsymbol{A}_1$ and $\boldsymbol{\mathcal{A}}_2 = \alpha \boldsymbol{I} + \boldsymbol{A}_2$ then, by Proposition 5, we know that $\lambda_1^{\boldsymbol{A}_1} = \lambda_1^{\boldsymbol{A}_2} = 1 - \alpha$. Therefore the first eigenvalues does not tell us which network converge faster to $\boldsymbol{\theta}^*$. We know, from the proof of Proposition 5 that $LT(\epsilon, \mathcal{G})$ depend on eigenvalues of $\boldsymbol{A}$ and not $\boldsymbol{\mathcal{A}}$. Thus, since $\lambda_1^{\boldsymbol{A}_1} = \lambda_1^{\boldsymbol{A}_2} = 1 - \alpha$ we can say that

$$LT(\epsilon, \mathcal{G}_1) \leq LT(\epsilon, \mathcal{G}_2) \iff \lambda_2^{\boldsymbol{A}_1} \leq \lambda_2^{\boldsymbol{A}_2} \tag{3.30}$$

Lemma 2 (Horn and Johnson, 1985) *Given two commuting matrix $\boldsymbol{C}$ and $\boldsymbol{D}$, there exists a unitary matrix $\boldsymbol{U}$ such that $\boldsymbol{U}^{-1}\boldsymbol{C}\boldsymbol{U} = \Lambda^{\boldsymbol{C}}$ and $\boldsymbol{U}^{-1}\boldsymbol{D}\boldsymbol{U} = \Lambda^{\boldsymbol{D}}$. where $\Lambda^{\boldsymbol{C}}, \Lambda^{\boldsymbol{D}}$ are diagonal matrices with eigenvalues as elements. Thus we get*

$$\boldsymbol{C} + \boldsymbol{D} = \boldsymbol{U}(\Lambda^{\boldsymbol{C}} + \lambda^{\boldsymbol{D}})\boldsymbol{U}^{-1}$$

Applying this lemma to both $\boldsymbol{\mathcal{A}}_1$ and $\boldsymbol{\mathcal{A}}_2$ and considering only the second largest eigenvalue, we obtain that

$$\lambda_2^{\boldsymbol{\mathcal{A}}_1} = \alpha + \lambda_2^{\boldsymbol{A}_1},\ \lambda_2^{\boldsymbol{\mathcal{A}}_2} = \alpha + \lambda_2^{\boldsymbol{A}_2}$$

Therefore if

$$\lambda_2^{\boldsymbol{\mathcal{A}}_1} \leq \lambda_2^{\boldsymbol{\mathcal{A}}_2} \iff \lambda_2^{\boldsymbol{A}_1} \leq \lambda_2^{\boldsymbol{A}_2} \tag{3.31}$$

Using (3.29), (3.30) and (3.31) togheter we get exactly the result of Corollary 5.1.

□

Appendix B

Supplementary Materials

Convergence toward the Truth

Let us consider a network with a malevolent stubborn agent, the “spreader”, s who affects the opinion dynamics declaring $\theta_s \neq \theta^*$ and a benevolent policymaker p (still stubborn) who want citizens to be as more informed as possible and thus to minimize the distance of steady state opinion vectors to the truth[35]

$$u_p(\boldsymbol{\mu}) = -(\boldsymbol{\mu} - \mathbf{1}\theta^*)^2 \tag{3.32}$$

Since it is not always possible to directly affect the network structure, we wonder if there is an action (declaration of θ_p) that the policymaker p can do to promote the spread of truth against the presence of stubborn agents in the society. It is clear that (30) is maximized when all agents have opinions equal to θ^*. In the next proposition, we show that if the influence of the two stubborn (p=policymaker and s=spreader) is not symmetric, then it is not possible to reach exactly the truth

Proposition 7 *If the two stubborn agents have different influence over agents,* $\boldsymbol{a}_p \neq \boldsymbol{a}_s \neq 0$*, and* $\theta_s \neq \theta^*$ *then*

$$\nexists\ \theta_g\ :\ \boldsymbol{\mu} = \boldsymbol{\theta^*}$$

If stubborn agents have equal influence over agents, $\boldsymbol{a}_p = \boldsymbol{a}_s$*, and one among them declare*

$$\theta_p = 2\theta^* - \theta_s$$

[35] Notice that at this stage we assume the stubborn agent to be naive, namely the opinion $\theta_s \neq \theta^*$ is not the result of a maximization process, but it is exogenous.

then all agents in the society learn the truth:

$$\boldsymbol{\mu} = \boldsymbol{\theta}^*$$

Notice that this results can be extended to a society with more than two stubborn, in the appendix we provide the characterization of the result with many stubborn.

Even if, whenever $\boldsymbol{a}_p \neq \boldsymbol{a}_s$ it is not possible for a policymaker to chose θ_p such that $\boldsymbol{\mu} = \boldsymbol{\theta}^*$ it is still possible to get close enough to the truth by declaring an opinion that minimizes the distance with the steady state opinions' vector.

Proposition 8 *If all agents have normal prior distribution at $t = 0$ and "imperfect recall" and there a malevolent spreader s with fixed opinion $\theta_s \neq \theta^*$ and a benevolent policymaker p with utility function (3.32), then p would declare the following opinion*

$$\theta_p = \frac{\boldsymbol{1}'\boldsymbol{C}\boldsymbol{a}_p - (\boldsymbol{C}(\boldsymbol{D}(\boldsymbol{I}-\boldsymbol{G})\mathbf{1}))'\boldsymbol{a}_p}{\boldsymbol{a}_p'\boldsymbol{C}'\boldsymbol{C}\boldsymbol{a}_p}\theta^* - \frac{\boldsymbol{a}_p'\boldsymbol{C}'\boldsymbol{C}\boldsymbol{a}_s}{\boldsymbol{a}_p'\boldsymbol{C}'\boldsymbol{C}\boldsymbol{a}_p}\theta_s$$

then steady state vector of beliefs (opinions) is

$$\boldsymbol{\mu} = \boldsymbol{C}\left(\left(\boldsymbol{D}(\boldsymbol{I}-\boldsymbol{G})\mathbf{1} + \boldsymbol{a}_p\frac{\boldsymbol{1}'\boldsymbol{C}\boldsymbol{a}_p - (\boldsymbol{C}(\boldsymbol{D}(\boldsymbol{I}-\boldsymbol{G})\mathbf{1}))'\boldsymbol{a}_p}{\boldsymbol{a}_p'\boldsymbol{C}'\boldsymbol{C}\boldsymbol{a}_p}\right)\theta^* + \left(\boldsymbol{a}_s - \boldsymbol{a}_p\frac{\boldsymbol{a}_p'\boldsymbol{C}'\boldsymbol{C}\boldsymbol{a}_s}{\boldsymbol{a}_p'\boldsymbol{C}'\boldsymbol{C}\boldsymbol{a}_p}\right)\theta_s\right) \tag{3.33}$$

If the stubborn is sophisticated as in Section 3.3, the policy maker has to keep that into account. To have more tractable results we assume, without loss of generality, that $\theta_p = \theta^* = 0$

Proposition 9 *If all agents have normal prior distribution at $t = 0$ and "imperfect recall" and there is a sophisticated malevolent spreader s that solves the problem in (3.12) and a benevolent policymaker p with utility function (3.32) where $\theta_p = \theta^* = 0$, then the declared opinions are*

$$\begin{cases} \theta_s^d = \frac{\boldsymbol{1}'\boldsymbol{C}\boldsymbol{a}_s(\boldsymbol{a}_p'\boldsymbol{C}'\boldsymbol{C}\boldsymbol{a}_p+k)}{j}\theta_s \\ \theta_p^d = -\frac{\boldsymbol{a}_p'\boldsymbol{C}'\boldsymbol{C}\boldsymbol{a}_s}{\boldsymbol{a}_p'\boldsymbol{C}'\boldsymbol{C}\boldsymbol{a}_p+k}\frac{\boldsymbol{1}'\boldsymbol{C}\boldsymbol{a}_s(\boldsymbol{a}_p'\boldsymbol{C}'\boldsymbol{C}\boldsymbol{a}_p+k)}{j}\theta_s \end{cases} \tag{3.34}$$

where $j = k(k+\boldsymbol{a}_s'\boldsymbol{C}'\boldsymbol{C}\boldsymbol{a}_s+\boldsymbol{a}_p'\boldsymbol{C}'\boldsymbol{C}\boldsymbol{a}_p)+\boldsymbol{a}_s'\boldsymbol{C}'\boldsymbol{C}\boldsymbol{a}_s\boldsymbol{a}_p'\boldsymbol{C}'\boldsymbol{C}\boldsymbol{a}_p-\boldsymbol{a}_s'\boldsymbol{C}'\boldsymbol{C}\boldsymbol{a}_p\boldsymbol{a}_p'\boldsymbol{C}'\boldsymbol{C}\boldsymbol{a}_s$ is a constant.

Then steady state vector of beliefs (opinions) is

$$\boldsymbol{\mu} = \boldsymbol{C}\left(\frac{\boldsymbol{1}'\boldsymbol{C}\boldsymbol{a}_s(\boldsymbol{a}_p'\boldsymbol{C}'\boldsymbol{C}\boldsymbol{a}_p + k)}{j}\left(\boldsymbol{a}_s - \boldsymbol{a}_p\frac{\boldsymbol{a}_p'\boldsymbol{C}'\boldsymbol{C}\boldsymbol{a}_s}{\boldsymbol{a}_p'\boldsymbol{C}'\boldsymbol{C}\boldsymbol{a}_p + k}\right)\right)\theta_s \tag{3.35}$$

We now provide proofs of proposition in this section.

Proof Proposition 7

Notice that since $\boldsymbol{\mathcal{A}}$ is stochastic matrix then $(\boldsymbol{I} - \boldsymbol{A})\mathbf{1} = \boldsymbol{D}\mathbf{1} + \boldsymbol{a}_{s1} + \boldsymbol{a}_{s2}$, then

- If $\boldsymbol{a}_{s1} = \boldsymbol{a}_{s2} = 0$

  $$\boldsymbol{\mu} = (\boldsymbol{I} - \boldsymbol{DG} - \boldsymbol{A})^{-1}(\boldsymbol{D}(\boldsymbol{I} - \boldsymbol{G})\boldsymbol{\theta}^*)$$

  where $(\boldsymbol{I} - \boldsymbol{DG} - \boldsymbol{A})\mathbf{1} = (\boldsymbol{D}(\boldsymbol{I} - \boldsymbol{G})\mathbf{1})$, thus

  $$\boldsymbol{\mu} = (\boldsymbol{I} - \boldsymbol{DG} - \boldsymbol{A})^{-1}(\boldsymbol{D}(\boldsymbol{I} - \boldsymbol{G})\mathbf{1})\,\theta^*$$

  $$\Rightarrow \boldsymbol{\mu} = \underbrace{(\boldsymbol{D} - \boldsymbol{DG})^{-1}(\boldsymbol{D} - \boldsymbol{DG})}_{\boldsymbol{I}}\mathbf{1}\theta^* = \boldsymbol{\theta}^*$$

- If $\boldsymbol{\theta}_{s1} = \boldsymbol{\theta}_{s2} = \theta^*$

  $$\boldsymbol{\mu} = (\boldsymbol{I} - \boldsymbol{DG} - \boldsymbol{A})^{-1}(\boldsymbol{D}(\boldsymbol{I} - \boldsymbol{G})\boldsymbol{\theta}^* + \boldsymbol{a}_{s1}\theta^* + \boldsymbol{a}_{s2}\theta^*)$$

  where $(\boldsymbol{I} - \boldsymbol{DG} - \boldsymbol{A})\mathbf{1} = (\boldsymbol{D}(\boldsymbol{I} - \boldsymbol{G})\mathbf{1} + \boldsymbol{a}_{s1} + \boldsymbol{a}_{s2})$, thus

  $$\boldsymbol{\mu} = (\boldsymbol{D} + \boldsymbol{a}_{s1} + \boldsymbol{a}_{s2} - \boldsymbol{DG})^{-1}(\boldsymbol{D}(\boldsymbol{I} - \boldsymbol{G})\mathbf{1} + \boldsymbol{a}_{s1} + \boldsymbol{a}_{s2})\,\theta^*$$

  $$\Rightarrow \boldsymbol{\mu} = \underbrace{(\boldsymbol{D}(\boldsymbol{I} - \boldsymbol{G}) + \boldsymbol{a}_{s1} + \boldsymbol{a}_{s2})^{-1}((\boldsymbol{D}(\boldsymbol{I} - \boldsymbol{G}) + \boldsymbol{a}_{s1} + \boldsymbol{a}_{s2})}_{\mathbf{1}}\boldsymbol{\theta}^* = \theta^*$$

- $\boldsymbol{\mu} = \boldsymbol{\theta}^*$ if

  $$\boldsymbol{a}_{s1}\theta_{s1} + \boldsymbol{a}_{s2}\theta_{s2} = (\boldsymbol{a}_{s1} + \boldsymbol{a}_{s2})\,\theta^*$$

  $$\boldsymbol{a}_{s1}\theta_{s1} = (\boldsymbol{a}_{s1} + \boldsymbol{a}_{s2})\,\theta^* - \boldsymbol{a}_{s2}\theta_{s2}$$

  $$\boldsymbol{a}_{s1}\theta_{s1} = \boldsymbol{a}_{s1}\theta^* + \boldsymbol{a}_{s2}(\theta^* - \theta_{s2})$$

If θ_{s1} it is a scalar, the system has only one solution

$$\theta_{s1} = 2\theta^* - \theta_{s2}$$

if and only if $\boldsymbol{a}_{s1} = \boldsymbol{a}_{s2}$.

On the other hand, if $\boldsymbol{a}_{s1} \neq \boldsymbol{a}_{s2}$ if θ_{s1} is a scalar is not possible for one of the two stubborn to compensate the distortion created by the other.

(Double check) If $\boldsymbol{a}_{s1} = \boldsymbol{a}_{s2} = \boldsymbol{a}_s$ and $\theta_{s1} = 2\theta^* - \theta_{s2}$

$$\boldsymbol{\mu} = (\boldsymbol{I} - \boldsymbol{DG} - \boldsymbol{A})^{-1}(\boldsymbol{D}(\boldsymbol{I} - \boldsymbol{G})\boldsymbol{\theta^*} + \boldsymbol{a}_{s1}(2\theta^* - \theta_{s2}) + \boldsymbol{a}_{s2}\theta_{s_2})$$

$$\boldsymbol{\mu} = (\boldsymbol{I} - \boldsymbol{DG} - \boldsymbol{A})^{-1}(\boldsymbol{D}(\boldsymbol{I} - \boldsymbol{G})\boldsymbol{\theta^*} + 2\boldsymbol{a}_s\theta^*)$$

$$\boldsymbol{\mu} = (\boldsymbol{I} - \boldsymbol{DG} - \boldsymbol{A})^{-1}(\boldsymbol{D}(\boldsymbol{I} - \boldsymbol{G})\mathbf{1} + 2\boldsymbol{a}_s)\theta^*$$

which is equivalent to

$$\boldsymbol{\mu} = (\boldsymbol{D} + 2\boldsymbol{a}_s - \boldsymbol{DG})^{-1}(\boldsymbol{D}(\boldsymbol{I} - \boldsymbol{G})\mathbf{1} + 2\boldsymbol{a}_s)\theta^*$$

The first two bullet points discuss the conditions under which the truth is always reached while the third bullet point prove Proposition 7. Now we generalize the result to more than 2 stubborn.

If the cardinality of the set of stubborn agents is S and $\boldsymbol{a}_{s1} = \boldsymbol{a}_{s2} = ... = \boldsymbol{a}_{ss} = \boldsymbol{a}_s$ then $\boldsymbol{\mu} = \boldsymbol{\theta^*}$ if

$$\boldsymbol{a}_s\theta_{s1} + \sum_{s=2}^{S} \boldsymbol{a}_s\theta_{ss} = \left(\sum_{s=1}^{S} \boldsymbol{a}_{ss}\right)\theta^*$$

$$\Rightarrow \boldsymbol{a}_s\theta_{s1} = \left(\boldsymbol{a}_s + \sum_{s=2}^{S} \boldsymbol{a}_{ss}\right)\theta^* - \sum_{s=2}^{S} \boldsymbol{a}_{ss}\theta_{ss}$$

$$\Rightarrow \boldsymbol{a}_s\theta_{s1} = \boldsymbol{a}_s\theta^* + \sum_{s=2}^{S} \boldsymbol{a}_{ss}(\theta^* - \theta_{ss})$$

$$\Rightarrow \theta_s = \theta^* + \sum_{s=2}^{S} \frac{\boldsymbol{a}_{ss}(\theta^* - \theta_{ss})}{\boldsymbol{a}_{s1}}$$

□

Proof Proposition 8

Maximizing the utility function (3.32)

$$\max_{\theta_p} -(\boldsymbol{\mu} - \mathbf{1}\theta^*)^2$$

$$\max_{\theta_p} -\boldsymbol{\mu}'\boldsymbol{\mu} + 2(\mathbf{1}\theta^*)'\boldsymbol{\mu}$$

since

$$\boldsymbol{\mu} = \boldsymbol{C}\left(\boldsymbol{D}(\boldsymbol{I} - \boldsymbol{G})\boldsymbol{\theta}^* + \boldsymbol{a}_p\theta_p + \boldsymbol{a}_s\theta_s\right)$$

substituting $\boldsymbol{\mu}$ in the problem and considering only elements depending on θ_p we get

$$\max_{\theta_p} \; -2\left(\boldsymbol{C}\left(\boldsymbol{D}(\boldsymbol{I} + \boldsymbol{G})\boldsymbol{\theta}^*\right)\right)'\boldsymbol{a}_p\theta_p - 2\boldsymbol{a}_p'\boldsymbol{C}'\boldsymbol{C}\boldsymbol{a}_s\theta_p\theta_s - \boldsymbol{a}_p'\boldsymbol{C}'\boldsymbol{C}\boldsymbol{a}_p\theta_p^2 + 2(\mathbf{1}\theta^*)'\boldsymbol{C}\boldsymbol{a}_p\theta_p$$

we now solve the First Order Condition

$$-2\left(\boldsymbol{C}\left(\boldsymbol{D}(\boldsymbol{I} - \boldsymbol{G})\boldsymbol{\theta}^*\right)\right)'\boldsymbol{a}_p - 2\boldsymbol{a}_p'\boldsymbol{C}'\boldsymbol{C}\boldsymbol{a}_s\theta_s - 2\boldsymbol{a}_p'\boldsymbol{C}'\boldsymbol{C}\boldsymbol{a}_p\theta_p + 2(\mathbf{1}\theta^*)'\boldsymbol{C}\boldsymbol{a}_p = 0$$

$$-\left(\boldsymbol{C}\left(\boldsymbol{D}(\boldsymbol{I} - \boldsymbol{G})\mathbf{1}\right)\right)'\boldsymbol{a}_p\theta^* - \boldsymbol{a}_p'\boldsymbol{C}'\boldsymbol{C}\boldsymbol{a}_s\theta_s - \boldsymbol{a}_p'\boldsymbol{C}'\boldsymbol{C}\boldsymbol{a}_p\theta_p + \mathbf{1}'\boldsymbol{C}\boldsymbol{a}_p\theta^* = 0$$

Thus the optimal policy maker's declaration is

$$\theta_p = \frac{\mathbf{1}'\boldsymbol{C}\boldsymbol{a}_p - \left(\boldsymbol{C}\left(\boldsymbol{D}(\boldsymbol{I} - \boldsymbol{G})\mathbf{1}\right)\right)'\boldsymbol{a}_p}{\boldsymbol{a}_p'\boldsymbol{C}'\boldsymbol{C}\boldsymbol{a}_p}\theta^* - \frac{\boldsymbol{a}_p'\boldsymbol{C}'\boldsymbol{C}\boldsymbol{a}_s}{\boldsymbol{a}_p'\boldsymbol{C}'\boldsymbol{C}\boldsymbol{a}_p}\theta_s$$

Substituting in $\boldsymbol{\mu}$ we get exactly

$$\boldsymbol{\mu} = \boldsymbol{C}\left(\boldsymbol{D}(\boldsymbol{I} - \boldsymbol{G})\boldsymbol{\theta}^* + \boldsymbol{a}_p\left(\frac{\mathbf{1}'\boldsymbol{C}\boldsymbol{a}_p - \left(\boldsymbol{C}\left(\boldsymbol{D}(\boldsymbol{I} + \boldsymbol{G})\mathbf{1}\right)\right)'\boldsymbol{a}_p}{\boldsymbol{a}_p'\boldsymbol{C}'\boldsymbol{C}\boldsymbol{a}_p}\theta^* - \frac{\boldsymbol{a}_p'\boldsymbol{C}'\boldsymbol{C}\boldsymbol{a}_s}{\boldsymbol{a}_p'\boldsymbol{C}'\boldsymbol{C}\boldsymbol{a}_p}\theta_s\right) + \boldsymbol{a}_s\theta_s\right)$$

$$\boldsymbol{\mu} = \boldsymbol{C}\left(\left(\boldsymbol{D}(\boldsymbol{I} - \boldsymbol{G})\mathbf{1} + \boldsymbol{a}_p\frac{\mathbf{1}'\boldsymbol{C}\boldsymbol{a}_p - \left(\boldsymbol{C}\left(\boldsymbol{D}(\boldsymbol{I} - \boldsymbol{G})\mathbf{1}\right)\right)'\boldsymbol{a}_p}{\boldsymbol{a}_p'\boldsymbol{C}'\boldsymbol{C}\boldsymbol{a}_p}\right)\theta^* + \left(\boldsymbol{a}_s - \boldsymbol{a}_p\frac{\boldsymbol{a}_p'\boldsymbol{C}'\boldsymbol{C}\boldsymbol{a}_s}{\boldsymbol{a}_p'\boldsymbol{C}'\boldsymbol{C}\boldsymbol{a}_p}\right)\theta_s\right) \tag{3.36}$$

□

Proof Proposition 9

There are two stubborn, one controlled by the policy maker p and spreader of misinformation s.

$$\max_{\theta_p^d} \; u_s(\boldsymbol{\mu}): \; -(\boldsymbol{\mu}-\mathbf{1}\theta_p)^2 - k(\theta^* - \theta_p^d)^2$$

We assume, without loss of generality, that $\theta_p = \theta^* = 0$

$$\max_{\theta_p^d} \; u_s(\boldsymbol{\mu}): \; -(\boldsymbol{\mu}-\mathbf{0})^2 - k(-\theta_p^d)^2$$

$$\boldsymbol{\mu} = \boldsymbol{C}\left(\boldsymbol{a}_p\theta_p^d + \boldsymbol{a}_s\theta_s^d\right) \tag{3.37}$$

$$\max_{\theta_p^d} \; -\boldsymbol{\mu}'\boldsymbol{\mu} - k(-\theta_p^d)^2$$

$$\max_{\theta_p^d} \; -2\boldsymbol{a}_p'\boldsymbol{C}'\boldsymbol{C}\boldsymbol{a}_s\theta_p^d\theta_s^d - \boldsymbol{a}_p'\boldsymbol{C}'\boldsymbol{C}\boldsymbol{a}_p\theta_p^{d2} - k(-\theta_p^d)^2$$

F.O.C.

$$-\boldsymbol{a}_p'\boldsymbol{C}'\boldsymbol{C}\boldsymbol{a}_s\theta_s^d - \boldsymbol{a}_p'\boldsymbol{C}'\boldsymbol{C}\boldsymbol{a}_p\theta_p^d - k\theta_p^d = 0$$

Solving for θ_p^d

$$\theta_p^d = -\frac{\boldsymbol{a}_p'\boldsymbol{C}'\boldsymbol{C}\boldsymbol{a}_s}{\boldsymbol{a}_p'\boldsymbol{C}'\boldsymbol{C}\boldsymbol{a}_p + k}\theta_s^d$$

by symmetry

$$\theta_s^d = \frac{\mathbf{1}'\boldsymbol{C}\boldsymbol{a}_s}{\boldsymbol{a}_s'\boldsymbol{C}'\boldsymbol{C}\boldsymbol{a}_s + k}\theta_s - \frac{\boldsymbol{a}_s'\boldsymbol{C}'\boldsymbol{C}\boldsymbol{a}_p}{\boldsymbol{a}_s'\boldsymbol{C}'\boldsymbol{C}\boldsymbol{a}_s + k}\theta_p^d$$

Substituting θ_p^d

$$\Rightarrow \theta_s^d = \frac{\mathbf{1}'\boldsymbol{C}\boldsymbol{a}_s}{\boldsymbol{a}_s'\boldsymbol{C}'\boldsymbol{C}\boldsymbol{a}_s + k}\theta_s + \frac{\boldsymbol{a}_s'\boldsymbol{C}'\boldsymbol{C}\boldsymbol{a}_p}{\boldsymbol{a}_s'\boldsymbol{C}'\boldsymbol{C}\boldsymbol{a}_s + k}\frac{\boldsymbol{a}_p'\boldsymbol{C}'\boldsymbol{C}\boldsymbol{a}_s}{\boldsymbol{a}_p'\boldsymbol{C}'\boldsymbol{C}\boldsymbol{a}_p + k}\theta_s^d$$

Solving for θ_s^d

$$\theta_s^d\frac{k(k + \boldsymbol{a}_s'\boldsymbol{C}'\boldsymbol{C}\boldsymbol{a}_s + \boldsymbol{a}_p'\boldsymbol{C}'\boldsymbol{C}\boldsymbol{a}_p) + \boldsymbol{a}_s'\boldsymbol{C}'\boldsymbol{C}\boldsymbol{a}_s\boldsymbol{a}_p'\boldsymbol{C}'\boldsymbol{C}\boldsymbol{a}_p - \boldsymbol{a}_s'\boldsymbol{C}'\boldsymbol{C}\boldsymbol{a}_p\boldsymbol{a}_p'\boldsymbol{C}'\boldsymbol{C}\boldsymbol{a}_s}{(\boldsymbol{a}_s'\boldsymbol{C}'\boldsymbol{C}\boldsymbol{a}_s + k)(\boldsymbol{a}_p'\boldsymbol{C}'\boldsymbol{C}\boldsymbol{a}_p + k)} = \frac{\mathbf{1}'\boldsymbol{C}\boldsymbol{a}_s}{\boldsymbol{a}_s'\boldsymbol{C}'\boldsymbol{C}\boldsymbol{a}_s + k}\theta_s$$

Therefore

$$\Rightarrow \theta_s^d = \frac{\mathbf{1}'\boldsymbol{C}\boldsymbol{a}_s(\boldsymbol{a}_p'\boldsymbol{C}'\boldsymbol{C}\boldsymbol{a}_p + k)}{j}\theta_s \tag{3.38}$$

where $j = k(k + \boldsymbol{a}_s'\boldsymbol{C}'\boldsymbol{C}\boldsymbol{a}_s + \boldsymbol{a}_p'\boldsymbol{C}'\boldsymbol{C}\boldsymbol{a}_p) + \boldsymbol{a}_s'\boldsymbol{C}'\boldsymbol{C}\boldsymbol{a}_s\boldsymbol{a}_p'\boldsymbol{C}'\boldsymbol{C}\boldsymbol{a}_p - \boldsymbol{a}_s'\boldsymbol{C}'\boldsymbol{C}\boldsymbol{a}_p\boldsymbol{a}_p'\boldsymbol{C}'\boldsymbol{C}\boldsymbol{a}_s$ is a constant.

$$\theta_p^d = -\frac{\boldsymbol{a}_p'\boldsymbol{C}'\boldsymbol{C}\boldsymbol{a}_s}{\boldsymbol{a}_p'\boldsymbol{C}'\boldsymbol{C}\boldsymbol{a}_p + k}\frac{\mathbf{1}'\boldsymbol{C}\boldsymbol{a}_s(\boldsymbol{a}_p'\boldsymbol{C}'\boldsymbol{C}\boldsymbol{a}_p + k)}{j}\theta_s \tag{3.39}$$

θ_s^d and θ_p^d are both decreasing in k.

Substituting (3.38) and (3.39) into (3.37) we obtain

$$\boldsymbol{\mu} = \boldsymbol{C}\left(-\boldsymbol{a}_p\frac{\boldsymbol{a}_p'\boldsymbol{C}'\boldsymbol{C}\boldsymbol{a}_s}{\boldsymbol{a}_p'\boldsymbol{C}'\boldsymbol{C}\boldsymbol{a}_p + k}\frac{\mathbf{1}'\boldsymbol{C}\boldsymbol{a}_s(\boldsymbol{a}_p'\boldsymbol{C}'\boldsymbol{C}\boldsymbol{a}_p + k)}{j} + \boldsymbol{a}_s\frac{\mathbf{1}'\boldsymbol{C}\boldsymbol{a}_s(\boldsymbol{a}_p'\boldsymbol{C}'\boldsymbol{C}\boldsymbol{a}_p + k)}{j}\right)\theta_s$$

$$\boldsymbol{\mu} = \boldsymbol{C}\left(\frac{\mathbf{1}'\boldsymbol{C}\boldsymbol{a}_s(\boldsymbol{a}_p'\boldsymbol{C}'\boldsymbol{C}\boldsymbol{a}_p + k)}{j}\left(\boldsymbol{a}_s - \boldsymbol{a}_p\frac{\boldsymbol{a}_p'\boldsymbol{C}'\boldsymbol{C}\boldsymbol{a}_s}{\boldsymbol{a}_p'\boldsymbol{C}'\boldsymbol{C}\boldsymbol{a}_p + k}\right)\right)\theta_s$$

Bibliography

Aarts, H. and Dijksterhuis, A. (2003). The silence of the library: environment, situational norm, and social behavior. *Journal of personality and social psychology*, 84(1):18. 9

Acemoglu, D., Dahleh, M. A., Lobel, I., and Ozdaglar, A. (2011). Bayesian learning in social networks. *The Review of Economic Studies*, 78(4):1201–1236. 92

Acemoglu, D. and Jackson, M. O. (2014). History, expectations, and leadership in the evolution of social norms. *The Review of Economic Studies*, 82(2):423–456. 13, 70

Acemoglu, D. and Ozdaglar, A. (2011). Opinion dynamics and learning in social networks. *Dynamic Games and Applications*, 1(1):3–49. 89

Agaev, R. and Chebotarev, P. (2005). On the spectra of nonsymmetric laplacian matrices. *Linear Algebra and its Applications*, 399:157–168. 119

Akerlof, G. (1976). The economics of caste and of the rat race and other woeful tales. *The Quarterly Journal of Economics*, pages 599–617. 9

Akerlof, G. A. and Dickens, W. T. (1982). The economic consequences of cognitive dissonance. *The American economic review*, 72(3):307–319. 13

Akerlof, G. A. and Kranton, R. E. (2000). Economics and identity. *The Quarterly Journal of Economics*, 115(3):715–753. 14

Alesina, A., Miano, A., and Stantcheva, S. (2018). Immigration and redistribution. Technical report, National Bureau of Economic Research. 52

Alger, I. and Weibull, J. W. (2013). Homo moralis—preference evolution under incomplete information and assortative matching. *Econometrica*, 81(6):2269–2302. 20

Bala, V. and Goyal, S. (1998). Learning from neighbours. *The Review of Economic Studies*, 65(3):595–621. 92

Ballester, C., Calvó-Armengol, A., and Zenou, Y. (2006). Who's who in networks. wanted: The key player. *Econometrica*, 74(5):1403–1417. 94

Bandura, A. (1993). Perceived self-efficacy in cognitive development and functioning. *Educational psychologist*, 28(2):117–148. 52, 57

Banerjee, A., Chandrasekhar, A. G., Duflo, E., and Jackson, M. O. (2013). The diffusion of microfinance. *Science*, 341(6144):1236498. 94

Battigalli, P. (2018). Game theory: Analysis of strategic thinking. 58, 59

Battigalli, P., Cerreia-Vioglio, S., Maccheroni, F., and Marinacci, M. (2015). Self-confirming equilibrium and model uncertainty. *The American Economic Review*, 105(2):646–77. 58

Battigalli, P., Gilli, M., and Molinari, M. C. (1992). *Learning and convergence to equilibrium in repeated strategic interactions: an introductory survey*. Università Commerciale" L. Bocconi", Istituto di Economia Politica. 59

Battigalli, P. and Guaitoli, D. (1988). Conjectural equilibrium. *mimeo*. 53, 58

Belsky, J., Robins, E., and Gamble, W. (1984). The determinants of parental competence. In *Beyond the dyad*, pages 251–279. Springer. 52, 55

Berry, J. W. (1997). Immigration, acculturation, and adaptation. *Applied Psychology*, 46(1):5–34. 29

Bessi, A., Coletto, M., Davidescu, G. A., Scala, A., Caldarelli, G., and Quattrociocchi, W. (2015). Science vs conspiracy: Collective narratives in the age of misinformation. *PloS One*, 10(2):e0118093. 90

Bimpikis, K., Ozdaglar, A., and Yildiz, E. (2016). Competitive targeted advertising over networks. *Operations Research*, 64(3):705–720. 93

Bindel, D., Kleinberg, J., and Oren, S. (2015). How bad is forming your own opinion? *Games and Economic Behavior*, 92:248–265. 93, 96

Bisin, A., Patacchini, E., Verdier, T., and Zenou, Y. (2011). Ethnic identity and labour market outcomes of immigrants in europe. *Economic Policy*, 26(65):57–92. 14

Bisin, A., Topa, G., and Verdier, T. (2004). Cooperation as a transmitted cultural trait. *Rationality and Society*, 16(4):477–507. 20

Bisin, A. and Verdier, T. (2001). The economics of cultural transmission and the dynamics of preferences. *Journal of Economic Theory*, 97(2):298–319. 11, 12, 13, 25, 51, 52, 54, 55, 57, 60, 61, 62, 63, 65, 73, 75

Bisin, A. and Verdier, T. (2012). The economics of cultural transmission and socialization. handbook of social economics, j. benhabib, a. bisin and mo jackson. 9, 13, 29

Bisin, A. and Verdier, T. (2017). On the joint evolution of culture and institutions. Technical report, National Bureau of Economic Research. 13

Bjorn, P. A. and Vuong, Q. H. (1984). Simultaneous equations models for dummy endogenous variables: a game theoretic formulation with an application to labor force participation. 24

Bloch, F. (2016). *Targeting and Pricing in Social Networks*, volume The Oxford Handbook of the Economics of Networks, 776-791. 93, 99

Bolletta, U. and Pin, P. (2018). Polarization through the endogenous network. 93, 96

Buechel, B., Hellmann, T., and Klößner, S. (2015). Opinion dynamics and wisdom under conformity. *Journal of Economic Dynamics and Control*, 52:240–257. 93

Bush, A. J., Martin, C. A., and Bush, V. D. (2004). Sports celebrity influence on the behavioral intentions of generation y. *Journal of Advertising Research*, 44(1):108–118. 89

Calabuig, V., Olcina, G., Panebianco, F., et al. (2014). The dynamics of personal norms and the emergence of cultural diversity. Technical report, University of Valencia, ERI-CES. 13

Candogan, O., Bimpikis, K., and Ozdaglar, A. (2012). Optimal pricing in networks with externalities. *Operations Research*, 60(4):883–905. 93

Carvalho, J.-P. (2016). Coordination and culture. *Economic Theory*, pages 1–27. 13, 20

Cavalli-Sforza, L. L. and Feldman, M. W. (1981). *Cultural transmission and evolution: a quantitative approach*. Number 16. Princeton University Press. 11, 13, 55

Cheeger, J. (1969). A lower bound for the smallest eigenvalue of the laplacian. In *Proceedings of the Princeton conference in honor of Professor S. Bochner*. 108

Cheung, M.-W. and Wu, J. (2018). On the probabilistic transmission of continuous cultural traits. *Journal of Economic Theory*, 174:300–323. 55, 73

Chung, F. R. (1996). Spectral graph theory (cbms regional conference series in mathematics, no. 92). 108, 109, 119, 120

Coleman, J. S., Katz, E., and Menzel, H. (1966). *Medical innovation: A Diffusion Study*. Bobbs-Merrill Co. 89

Coleman, P. K. and Karraker, K. H. (2000). Parenting self-efficacy among mothers of school-age children: Conceptualization, measurement, and correlates. *Family Relations*, 49(1):13–24. 52, 57, 60

Corazzini, L., Pavesi, F., Petrovich, B., and Stanca, L. (2012). Influential listeners: An experiment on persuasion bias in social networks. *European Economic Review*, 56(6):1276–1288. 95

Dalmazzo, A., Pin, P., and Scalise, D. (2014). Communities and social inefficiency with heterogeneous groups. *Journal of Economic Dynamics and Control*, 48:410–427. 13

Dasaratha, K., Golub, B., and Hak, N. (2018). Bayesian social learning in a dynamic environment. 93, 96

de Montigny, F. and Lacharité, C. (2005). Perceived parental efficacy: Concept analysis. *Journal of Advanced Nursing*, 49(4):387–396. 52, 57

DeGroot, M. H. (1974). Reaching a consensus. *Journal of the American Statistical Association*, 69(345):118–121. 14, 16, 91, 92

Del Vicario, M., Bessi, A., Zollo, F., Petroni, F., Scala, A., Caldarelli, G., Stanley, H. E., and Quattrociocchi, W. (2016). The spreading of misinformation online. *Proceedings of the National Academy of Sciences*, 113(3):554–559. 89

DeMarzo, P. M., Vayanos, D., and Zwiebel, J. (2003). Persuasion bias, social influence, and unidimensional opinions. *The Quarterly Journal of Economics*, 118(3):909–968. 14, 92

Donovan, W. L. and Leavitt, L. A. (1985). Simulating conditions of learned helplessness: The effects of interventions and attributions. *Child Development*, pages 594–603. 60

Donovan, W. L., Leavitt, L. A., and Walsh, R. O. (1997). Cognitive set and coping strategy affect mothers' sensitivity to infant cries: A signal detection approach. *Child Development*, 68(5):760–772. 60

Eshel, I., Samuelson, L., and Shaked, A. (1998). Altruists, egoists, and hooligans in a local interaction model. *American Economic Review*, pages 157–179. 21

Fainmesser, I. P. and Galeotti, A. (2015). Pricing network effects. *The Review of Economic Studies*, 83(1):165–198. 91, 93

Fainmesser, I. P. and Galeotti, A. (2016). Pricing network effects: Competition. 91, 93

Fainmesser, I. P. and Galeotti, A. (2018). The market for influence. 89

Feick, L. F. and Price, L. L. (1987). The market maven: A diffuser of marketplace information. *The Journal of Marketing*, pages 83–97. 89

Ferrara, E., Varol, O., Davis, C., Menczer, F., and Flammini, A. (2016). The rise of social bots. *Communications of the ACM*, 59(7):96–104. 90

Fudenberg, D. and Levine, D. K. (1993a). Self-confirming equilibrium. *Econometrica: Journal of the Econometric Society*, pages 523–545. 53, 58

Fudenberg, D. and Levine, D. K. (1993b). Steady state learning and nash equilibrium. *Econometrica: Journal of the Econometric Society*, pages 547–573. 59

Gale, D. and Kariv, S. (2003). Bayesian learning in social networks. *Games and Economic Behavior*, 45(2):329–346. 92

Galeotti, A. and Goyal, S. (2009). Influencing the influencers: a theory of strategic diffusion. *The RAND Journal of Economics*, 40(3):509–532. 91, 93

Godes, D. and Mayzlin, D. (2004). Using online conversations to study word-of-mouth communication. *Marketing Science*, 23(4):545–560. 89

Golub, B. and Jackson, M. O. (2010). Naive learning in social networks and the wisdom of crowds. *American Economic Journal: Microeconomics*, 2(1):112–149. 14, 91, 92, 101, 113

Golub, B. and Jackson, M. O. (2012). How homophily affects the speed of learning and best-response dynamics. *The Quarterly Journal of Economics*, 127(3):1287–1338. 14, 16, 91, 94, 101, 107, 109, 112, 113

Golub, B. and Sadler, E. (2016). *Learning in Social Networks*, volume Oxford Handbook on the Economics of Networks. Oxford: Oxford University Press. 92

Gottfried, J. and Shearer, E. (2016). News use across social media platforms 2016. *Pew Research Center*, 26. 89

Goyal, S., Heidari, H., and Kearns, M. (2014). Competitive contagion in networks. *Games and Economic Behavior*. 93

Grabisch, M., Mandel, A., Rusinowska, A., and Tanimura, E. (2017). Strategic influence in social networks. *Mathematics of Operations Research*. 93, 101

Gratton, G., Holden, R., and Kolotilin, A. (2017). When to drop a bombshell. *The Review of Economic Studies*, 85(4):2139–2172. 106

Grimm, V. and Mengel, F. (2014). An experiment on belief formation in networks. 95

Hayek, F. A. (1967). Notes on the evolution of systems of rules of conduct. *Studies in philosophy, politics and economics*, pages 66–81. 9

Horn, R. A. and Johnson, C. R. (1985). Matrix analysis cambridge university press. *New York*, 37. 121, 122

Jackson, A. P. and Huang, C. C. (2000). Parenting stress and behavior among single mothers of preschoolers: The mediating role of self-efficacy. *Journal of Social Service Research*, 26(4):29–42. 67

Jadbabaie, A., Molavi, P., Sandroni, A., and Tahbaz-Salehi, A. (2012). Non-bayesian social learning. *Games and Economic Behavior*, 76(1):210–225. 90, 93, 94, 96, 107, 114

Kartik, N. (2009). Strategic communication with lying costs. *The Review of Economic Studies*, 76(4):1359–1395. 104

Kartik, N., Ottaviani, M., and Squintani, F. (2007). Credulity, lies, and costly talk. *Journal of Economic theory*, 134(1):93–116. 104

Kessler, J. B. and Leider, S. (2012). Norms and contracting. *Management Science*, 58(1):62–77. 13, 20

Kimbrough, E. O. and Vostroknutov, A. (2016). Norms make preferences social. *Journal of the European Economic Association*, 14(3):608–638. 13, 20

Kooreman, P. (1994). Estimation of econometric models of some discrete games. *Journal of Applied Econometrics*, 9(3):255–268. 24

Kuhn, J. C. and Carter, A. S. (2006). Maternal self-efficacy and associated parenting cognitions among mothers of children with autism. *American Journal of Orthopsychiatry*, 76(4):564–575. 67

Kuran, T. and Sandholm, W. H. (2008). Cultural integration and its discontents. *The Review of Economic Studies*, 75(1):201–228. 10, 11, 13, 14, 16, 17, 25

López-Pérez, R. (2008). Aversion to norm-breaking: A model. *Games and Economic behavior*, 64(1):237–267. 13, 20

Lowes, S., Nunn, N., Robinson, J. A., and Weigel, J. L. (2017). The evolution of culture and institutions: Evidence from the kuba kingdom. *Econometrica*, 85(4):1065–1091. 34

Mandel, A. and Venel, X. (2017). Dynamic competition on social networks. 93

Martin, C. A. and Bush, A. J. (2000). Do role models influence teenagers' purchase intentions and behavior? *Journal of Consumer Marketing*, 17(5):441–453. 89

Mash, E. J. and Johnston, C. (1983). Parental perceptions of child behavior problems, parenting self-esteem, and mothers' reported stress in younger and older hyperactive and normal children. *Journal of Consulting and Clinical Psychology*, 51(1):86. 60

Michaeli, M. and Spiro, D. (2017). From peer pressure to biased norms. *American Economic Journal: Microeconomics*, 9(1):152–216. 13

Milgrom, P. and Roberts, J. (1991). Adaptive and sophisticated learning in normal form games. *Games and Economic Behavior*, 3(1):82–100. 53, 59

Mobilia, M. (2003). Does a single zealot affect an infinite group of voters? *Physical Review Letters*, 91(2):028701. 93, 97

Mobilia, M., Petersen, A., and Redner, S. (2007). On the role of zealotry in the voter model. *Journal of Statistical Mechanics: Theory and Experiment*, 2007(08):P08029. 93, 97

Molavi, P., Tahbaz-Salehi, A., and Jadbabaie, A. (2018). A theory of non-bayesian social learning. *Econometrica*, 86(2):445–490. 93, 95

Morris, S. and Shin, H. S. (2002). Social value of public information. *The American Economic Review*, 92(5):1521–1534. 16, 93, 96

Mossel, E., Sly, A., and Tamuz, O. (2015). Strategic learning and the topology of social networks. *Econometrica*, 83(5):1755–1794. 92

Nteta, T. M. and Wallsten, K. J. (2012). Preaching to the choir? religious leaders and a merican opinion on immigration reform. *Social Science Quarterly*, 93(4):891–910. 70

Olcina, G., Panebianco, F., and Zenou, Y. (2017). Conformism, social norms and the dynamics of assimilation. 14, 93, 96

Oreskes, N. (2004). The scientific consensus on climate change. *Science*, 306(5702):1686–1686. 90

Panebianco, F. (2014). Socialization networks and the transmission of interethnic attitudes. *Journal of Economic Theory*, 150:583–610. 13

Prummer, A. and Siedlarek, J.-P. (2017). Community leaders and the preservation of cultural traits. *Journal of Economic Theory*, 168:143–176. 54, 70

Ratkiewicz, J., Conover, M., Meiss, M. R., Gonçalves, B., Flammini, A., and Menczer, F. (2011). Detecting and tracking political abuse in social media. *ICWSM*, 11:297–304. 90

Reingen, P. H., Foster, B. L., Brown, J. J., and Seidman, S. B. (1984). Brand congruence in interpersonal relations: A social network analysis. *Journal of Consumer Research*, 11(3):771–783. 89

Rosenberg, D., Solan, E., and Vieille, N. (2009). Informational externalities and emergence of consensus. *Games and Economic Behavior*, 66(2):979–994. 92

Ryder, A. G., Alden, L. E., and Paulhus, D. L. (2000). Is acculturation unidimensional or bidimensional? a head-to-head comparison in the prediction of personality, self-identity, and adjustment. *Journal of personality and social psychology*, 79(1):49. 29

Schwartz, S. H. (1977). Normative influences on altruism. *Advances in experimental social psychology*, 10:221–279. 9

Sherman, J. and Morrison, W. J. (1950). Adjustment of an inverse matrix corresponding to a change in one element of a given matrix. *The Annals of Mathematical Statistics*, 21(1):124–127. 117

Silverman, C. (2016). This analysis shows how viral fake election news stories outperformed real news on facebook. *Buzzfeed News*, 16. 90

Soetevent, A. R. and Kooreman, P. (2007). A discrete-choice model with social interactions: with an application to high school teen behavior. *Journal of Applied Econometrics*, 22(3):599–624. 24

Tabellini, G. (2008). The scope of cooperation: Values and incentives. *The Quarterly Journal of Economics*, 123(3):905–950. 13, 20

Unger, D. G. and Wandersman, L. P. (1985). Social support and adolescent mothers: Action research contributions to theory and application. *Journal of Social Issues*, 41(1):29–45. 60

Varol, O., Ferrara, E., Davis, C. A., Menczer, F., and Flammini, A. (2017). Online human-bot interactions: Detection, estimation, and characterization. *arXiv preprint arXiv:1703.03107*. 90

Vaughan, D. (2013). Nurture vs. nurture: Endogenous parental and peer effects and the transmission of culture. Technical report, Working Papers, Banco de México. 16

Verdier, T. and Zenou, Y. (2015). The role of cultural leaders in the transmission of preferences. *Economics Letters*, 136:158–161. 54, 70

Verdier, T. and Zenou, Y. (2018). Cultural leader and the dynamics of assimilation. *Journal of Economic Theory*, 175:374–414. 54, 70

Wells-Parker, E., Miller, D. I., and Topping, J. S. (1990). Development of control-of-outcome scales and self-efficacy scales for women in four life roles. *Journal of Personality Assessment*, 54(3-4):564–575. 60

Yildiz, E., Ozdaglar, A., Acemoglu, D., Saberi, A., and Scaglione, A. (2013). Binary opinion dynamics with stubborn agents. *ACM Transactions on Economics and Computation (TEAC)*, 1(4):19. 14, 93, 97, 101

Young, H. P. (1998). Social norms and economic welfare. *European Economic Review*, 42(3):821–830. 9

Zenou, Y. (2016). *Key Players*, volume Oxford Handbook on the Economics of Networks. Oxford: Oxford University Press. 94, 99