Alma Mater Studiorum – Università di Bologna

DOTTORATO DI RICERCA IN

Economia

Ciclo XXVIII

Settore Concorsuale di afferenza: 13/A1

Settore Scientifico disciplinare: SECS P/01

SOCIAL INFLUENCE IN NETWORKS

Presentata da: **Ugo Bolletta**

Coordinatore Dottorato **Relatore**

Matteo Cervellati **Giulio Zanella**

Esame finale anno 2016

Contents

1 Polarization through the endogenous network **9**
1.1 Introduction . . . 10
1.2 Literature Review . . . 12
1.3 The model . . . 14
1.3.1 Network as given . . . 14
1.3.2 Endogenous network . . . 18
1.4 Dynamic setting . . . 20
1.4.1 Simulations . . . 22
1.5 Discussion . . . 25

2 Peer effects in schools **27**
2.1 Introduction . . . 28
2.2 Literature Review . . . 30
2.3 The Model . . . 32
2.3.1 Notation . . . 32
2.3.2 Payoffs and Preliminary Results . . . 32
2.3.3 Endogenous Network . . . 34
2.4 Applications . . . 37
2.4.1 Two Cultures: a Majority and a Minority . . . 38
2.4.2 Peer Effects in the Classroom . . . 40
2.4.3 Nash Equilibrium and Bonacich Centrality . . . 44
2.5 Discussion . . . 45

3 Search over the network **47**
3.1 Introduction . . . 48
3.2 Literature review . . . 49
3.3 The Model . . . 50
3.4 Search over the Network . . . 51

3.4.1 Expected degree 52
3.5 A routine for the meeting process 54

Appendices **63**

A Polarization through the endogenous network **65**
A.1 Payoff in equilibrium 65
A.2 Proof of Lemma 1 65
A.3 Payoff differential 66
A.4 Proof of Proposition 1 67
A.5 Proof of Proposition 2 69

B Peer effects in schools: a model **71**
B.1 Proof of Lemma 8 71
B.2 Proof of Proposition 9 71
B.3 Proof of Remark 10 73
B.4 Proof of Proposition 13 74
B.5 Proof of Proposition 14 76
B.6 Proof of Remark 19 77
B.7 Example - Three agents 77
B.8 Alternative utility function 79

C Search over the network **83**
C.1 Expected Degree 83
C.2 Stopping time 84

Introduction

In traditional economic analysis outcomes are often associated to a given social structure. In particular they were often analyzed situations where they exist either two-by-two interactions or global interactions. The study of networks focuses on all the other possible intermediate cases. Research during the past 30 years brought in an outstanding amount of facts documenting that the social structure is key in understanding real world phenomena, especially when strategic interaction occurs. Yet, lot has to be done to have a better understanding of the link between behavioral outcomes and the social structure, i.e. the network. This dependence of behaviors to other individuals choices is broadly defined as social influence.

In this thesis I focus in particular on this issue. My interest in research so far can be summarized by the two following points: *i)* understanding the relation between network formation and behavioral outcome, and *ii)* exploring the spillovers on multi-layered networks. In particular the first point is addressed through the first two chapters. The first studies a model that introduces a micro-foundation of processes of polarization. The general attitude of individuals to be associated with similar minded peers is a motive that may lead to segregation. Thus, when this happen we can observe that multiple social norms are adopted, although individuals share the same environmental and socio-political features. The second chapter is a more specific application of the previous findings. It is investigated more deeply the question about peer effects in the classroom. It is a matter of interest because through a deep knowledge of processes of peer effects we may be able to design optimally the composition of classrooms to foster positive influence among students. This would improve the overall achievement in the school, importantly keeping the level of resources fixed.

Finally the third chapter departs from the first two in terms of methodology, but it relates to the issue of multiplexing in network. By multiplexing we mean that there are more layers of a networks, with links not necessarily overlapping. The matter of interest in this context is what happens on a layer as a consequence of something happened on a different one. The application explored here is real and virtual networks. The paper in this topic is focused on a search process where agents want to find other agent. However they do not know on which platform nodes to be found are searching on, so they have to coordinate on a platform. This is one of the first attempts on this new frontier of research on social networks, and lot has to be done in this direction yet.

Through this thesis I aim to contribute to the existing literature in several aspects. The first chapter extends the current knowledge on polarization processes, as it proposes a micro-foundation for observed patterns. Among the main implications is that it is crucial to distinguish between transitional and long-run polarization. In the former case we can observe polarization even if the process converges to a consensus in the long-run. The latter will exhibit a reinforcement of polarization over time, reaching a maximum when partial consensus happens.

This is relevant for several issues. Examples are provided by polarization of political opinions in the U.S., which is documented both in the public, and in their representatives, i.e. in the Congress and in Senate. A different kind of field of application may be integrations of cultures. In this sense this paper stresses the importance of social interactions to foster integration, and also helps understanding long-run processes, for which we have only short evidence. Moreover, this chapters brings important theoretical contributions. The literature on network games focuses on how equilibrium outcomes depend on the network structures, which is exogenous. Literature on strategic network formation takes outcomes as given, and study network configuration that may arise and stability issues. In the first and second chapter I do both, merging this two streams of literature.

The insights are quite relevant as the implications on outcomes brought in by the endogeneity of the network are non-negligible. The importance of that is stressed more in the specific through a real world application of the model, which analyzes peer effects in the classroom. As we know in econometrics we observe correlated outcomes, but we can hardly distinguish whether this correlation depends on being connected, meaning that individuals influence each other given social connections, or whether correlation in behaviors is just a reason for individuals to form links. Having a proper model for that is relevant because allows us to focus on the channels of the process, and also have a deeper understanding of unexplained real world facts.

The final chapter is still within the range of issues covered by the topic of social influence. However it differs substantially from the first two. While we have so far described processes of influence between agents, we now move to study situation where the activity of an individual into a layer of the network influences her own activity onto a different layer of the same network. This case is analyzed in the framework of real and virtual networks. These are two layers of an agent network. The social contacts in real life resemble the contacts in virtual life. But clearly either individuals are on-line, or they are off-line. Therefore their activity on one network influence the other. It is discussed a network formation stage here too, but I explore a stochastic process. One of the reasons to do so is to give the due importance also to the randomness in meetings. Agents will keep some control over links, since their preferences still matter, but their choices are conditional on the meetings that took place, which follows a probabilistic scheme.

The main point of this chapter is to shed lights on intensity of usage of virtual social networks. In principle this work is part of a larger research agenda of mine, as I intend to further investigate this issue, both theoretically and empirically. The main insights so far are that individuals with strong preferences for meeting similar-minded friends, that face a static environment will likely spend lot of their time on virtual social networks, sacrificing their real side of life. This is a matter of interest that goes beyond economics, although it is a behavioral feature with potential influence on economic outcomes too.

The thesis is organized as follows. Each chapter is presented separately, but we collected all the bibliography. Proofs and calculations behind the results are in the appendix, that is organized by chapters. The section Conclusions collects some broader thoughts and some punchline on the work that is presented below.

Chapter 1

Polarization through the endogenous network[1]

Processes of polarization have been documented in several applications. Nevertheless most of the theories built so far show how herding behavior and convergence of opinions tend to be a regularity in several contexts. In this paper we develop a model where agents correct their heterogeneous initial opinions averaging the opinions of their neighbors. The key contribution is to let the network take place endogenously. While the most known results are derived assuming the network to be strongly connected, we show how this component depends on the initial distribution of opinions. To do so, we characterize the process letting naive learning be a best reply function for agents. This allows to study the incentives in linking choices on the primitive process. Results show that, if opinions are not distributed uniformly, there always exist conditions on the strength of the social influence to prevent the network to be connected. This causes polarization both in the transition and in the long run.

[1]This chapter is joint with Paolo Pin (Bocconi University).

1.1 Introduction

Understanding the processes of opinion formation is a matter of interest for several reasons. In a society different social norms can coexist in contexts that share lot of cultural and environmental features. Or similarly, political opinions tend to be concentrated around two different poles. It has been documented also that in the latter specific case, looking at the process dynamically we observe a reinforcement, and therefore opinions become more polarized over time.[2]

Our approach is to explain this issue through strategic network formation. The idea is that individuals are on one hand influenced by other opinions, but the choice of social connections, i.e. sources of opinions, is not arbitrary. Thus whether a society ends up having convergent opinions would strongly depend by the way individuals interact with each other. Clearly if a society is split and there is no interaction between the several groups, the evolution of social norm within subgroups is independent from the others, and therefore is likely to exhibit disagreement.

We assume that evolution of opinions follows a process where agents are boundedly rational, and then focus on structural conditions on the nature of social relations. Thus we study a process known as *naive learning*, and we contribute to the existing stream of literature endowing agents with additional rationality with respect to the fully myopic scenario, letting them form the network optimally. Our approach could be described through the formulation of the DeGroot learning dynamic (see DeGroot (1974)), which simply states that at each point in time agents have an opinion equal to the average of opinions in their neighborhood in the previous period. Formally denoting with x_i the opinion of an agent i, and with μ_i the average of the opinions of i's neighborhood we have

$$x_{i,s} = \mu_{i,s-1}$$

This process has been studied in several papers, of which we mention among others Hegselmann and Krause (2002), DeMarzo et al. (2003) and Golub and Jackson (2010). In particular the latter shows that in order to have convergence some conditions on the structure of the adjacency matrix are needed and, specifically, it has to be irreducible and aperiodic. The former translates in a network being strongly connected[3], while the latter

[2] See Pew Research Center, June, 2014, "Political Polarization in the American Public" for a study on public in US, Andris et al. (2015) for evidence from the US House of Representatives, and how polarization is in this case a dynamic process.

[3] See Brualdi and Ryser (1991) for a formal proof.

is easily satisfied assuming agents posing positive weight on their opinions.[4]

On top of these results, we let agents weight positively their own opinions, focusing therefore on conditions for which the network will be endogenously connected or disconnected. Thus the process that is going to be examined in this paper is

$$x_{i,s} = f\mu_i + (1-f)t_{i,s} \tag{1.1}$$

where f is a weight that we call flexibility, and we let $t_{i,s} = x_{i,s-1}$. Through that lot of emphasis on the first period, where agents have no connections and are endowed exogenously with an opinion. Hence we have $t_{i,1} = t_{i,0}$, since $x_{i,0} = t_{i,0}$. Thus we let agents weight their initial opinions. When the process is initialized the network is not yet formed, and then afterwards we want it to update according to the evolution of opinions. Formally the process embeds a strong inertia, because agents update the network according to the opinion exhibited in the previous period and the opinion of new neighbors. The two are clearly correlated. Nevertheless, if the original DeGroot process is taken into account, there will be no evolution of the network, and therefore we need this modification. Moreover, this is the core of naive learning, because agents reinforce their bias over time. In other words, they are not able to exploit all the information that is available.

We initialize the problem endowing agents with an opinion, and depending on that they will form the network optimally. To achieve that we identify a payoff structure, through reverse engineering, such that the averaging of opinions described by equation 1.1 is a best reply function of the game. This allows for an analysis of strategic network formation, with opinions evolving accordingly.

Through this small change in the structure of the problem we are able to translate the conditions on the connectedness of the network assumed in existing works, into conditions on the ex-ante spectrum of opinions. This is a matter of interest because on top of this process the literature may advance inquiring the possible sources of interference that could either amplify or weaken the polarization process.

In term of results we first identify the possible equilibria of the network formation stage. Given a distribution the process boils down to a unique equilibrium characterized by agents ordered on the opinion space, forming connections with adjacent agents. Therefore the network in equilibrium exhibits *homophily*, which is in line with broad evidence on

[4] The condition requires that the largest common divisor of the path length is 1. See Jackson (2008) for a discussion of this issue.

social networks, and more specifically with the literature on political opinions also, as documented by Gentzkow and Shapiro (2010), among others.

Hence we derive conditions on the initial distribution of opinions, sampled from an arbitrary distribution, such that in equilibrium the network will be either connected or disconnected. In particular we detect conditions on how profitable are connections, and how sensible are agents to others' opinions so that, when the distribution of agents is such that there is a mass of opinions that is relatively less represented than in its neighborhoods, the network will be disconnected. This leads to persistent divergence of opinion into a society. We believe it is an important contribution, especially considering that opinions take initially place under the influence of external factors. This could be the case in political opinions when considering media and politicians. Additional considerations could be drawn letting these external entities act strategically with respect to the network. That is another fundamental question that we leave for future work.

1.2 Literature Review

The literature has tried to explain disagreement using several techniques, clearly showing different possible outcomes. We want to contribute to the understanding of this issue using a network approach. This is not new, since research in networks has focused substantially on these issues. We could broadly define two approaches within this field, the *bayesian learning* and *naive learning*.

The former assumes fully rational agents that are thus able to exploit at best the information in their possess. In this context herding behavior is a natural result, and it is hard to escape from convergence of opinions. Therefore we should depart from full rationality. In this context researchers have been focusing on minimal conditions on bounded rationality such that non-convergence could be achieved as a result. One example of that is given by Yildiz et al. (2013) which assumes the presence of *stubborn agents* which do not pay attention to the information available, but simply stick to their own opinion. Our approach differ from this stream of literature since we assume a boundedly rational process, and add rationality on top of it.

Indeed this paper is closely related to the literature on *naive learning*, which began with the paper by DeGroot (1974). More recently this framework has been investigated in several other works that are a fortiori related to this paper.

DeMarzo et al. (2003) brings explicitly the network into the process of updating, and assumes an environment with *persuasion bias* and *uni-dimensional opinions*. The former means that agents hear the news, and then are influenced according to an exogenously given *listening matrix*, without accounting for repetition of information. This pushes toward convergence since the sources of information are over-counted because of the social structure brought in by the network. Uni-dimensional opinions instead means that opinions can be summarized by a single opinion. Although this is a result in that paper, we use here opinions that are uni-dimensional and can thus be summarized on an interval of a line.

Golub and Jackson (2010) exploiting some of the results in the paper summarized above, derive general conditions on the structure of the adjacency matrix such that consensus may occur. One of the main results, among others, is that the network should be strongly connected and its adjacency matrix be aperiodic, which means that the highest common divisor of the cycles length must be one. This is easily satisfied allowing for self cycles, with agent then taking into account their own opinion. Thus we keep this assumption but we study conditions on the distribution of opinions such that the network will be endogenously disconnected, which would then prevent consensus to happen. To the best of our knowledge, no other works attempts to do so.

With a completely different approach, Krause (2000) and Hegselmann and Krause (2002) do not analyze directly the network, but assume that the hearing matrix, which is in fact a stochastic matrix that summarizes a Markov process, is such that agents exhibit *bounded confidence*. This is implemented through an exogenous rule for the opinions to be taken into account only if they are similar enough. Under this circumstances, if this distance parameter is too relevant, consensus is not achieved. In our paper such a rule is not exogenous but is a consequence of the endogenous network, although follows from the payoff structure. The results are richer because they take into account relevant parameters to analyze different scenarios, and are derived for arbitrary distributions of opinions.

Another closely related paper to ours is Melguizo (2015) where the evolutions of opinions is based on *salience of attributes*, meaning the difference in behavior between individuals sharing a characteristic or lacking it. Agents are endowed with a vector of characteristics, and if there is a unique most salient characteristic the agents will assign, dynamically, growing weight to those who share the same trait. Since characteristics are binary groups can at most be two, and that is the case in which disagreement occurs. While this can be reasonable in several contexts, we believe the modeling explicitly the strategic network formation allows for a better understanding of the drivers of such results, and to derive richer results, too.

Most of the analysis of the model is here focused on the first period of the dynamic process, since this will determine the full evolution of opinions in the society. Nevertheless we contribute also to the existing literature on the speed of convergence to consensus. In this stream of literature we find Golub and Jackson (2012), among others. They study a process with agents belonging to a finite set of groups, and assume that interactions are more frequent toward same-type agents. In such a context convergence will always be reached, and thus it is shown how the stronger the same-group interactions (*homophily*), the slower the convergence to consensus. Here we show how, on top of this result, the speed of convergence is non linear due to the endogeneity of the network. In particular we identify thresholds on the diameter of the network such that the process speeds up dramatically and converge. Thus more homophilous networks tend to exhibit slower speed of convergence, but because of the endogenous network they will show also a more prominent non-linearity in the latest stages.

For different reasons, this paper is also related on the literature of network games with endogenous networks. On of the first attempts in this context is provided by Galeotti and Goyal (2010), and followed by Kinateder and Merlino (Kinateder and Merlino). These papers examine games of public goods and thus they differ substantially from the game proposed here, which formally is a game of complements. A similar model to the one in this paper is analyzed in Bolletta (2015). The main difference is that here we study directed networks, following therefore a fully non-cooperative approach. In that paper instead the model is solved under *pairwise Nash stability*, and the undirected network allows to focus on the agents less prone to form links. Under a perspective of optimal policy design that aim to foster interactions, the analysis identifies individuals more susceptible to incentives.

The rest of the paper is organized as follows. Next section describes the model, analyzed in steps assuming exogenous network first in section 1.3.1, then letting it be endogenous in section 1.3.2. The latter contains the main results of the paper. Thus we analyze the full dynamic process and speed of convergence in section 1.4. Discussion in section 2.5 concludes. Proofs are in Appendix.

1.3 The model

1.3.1 Network as given

Consider the following one-shot game between two players.[5] Each player i is characterized by an *opinion* (or *type*) $t_i \in [0, 1]$ and by a *flexibility* $f \in [0, 1]$. The action of each player

[5] The two players version of this game is similar to Bisin et al. (2006).

i is $x_i \in \mathbb{R}$, and the payoff for agent i is, given the action x_j of the other player,

$$\pi_i(x_i, x_j) = V - f(x_i - x_j)^2 - (1-f)(x_i - t_i)^2 \quad , \tag{1.2}$$

V is the value from the interaction, and then there are some costs for adaptation and coordination. V here could represent the idea that other agents have information that is useful, and we assume that it is linear and symmetric across agents.[6] The parameter f weights these two components, and that is why we have called it flexibility. One can already see that agents with low flexibility weigh more adaptation. High flexibility agents weigh more the coordination term.

Now consider a directed network with n agents, where the neighbors of node i are given by the set d_i, with cardinality $k_i = |d_i|$. We let the network be directed for two main reasons. In terms of interpretation we believe that in opinion formation, where the structure of social interactions has been often represented through a "listening matrix" is more consistent with what we observe. Moreover there is a technical reason, that is we solve the model through a fully non-cooperative approach.[7] The same game as above is played on the network but now an agent must choose the same action as before, taking into account all her neighbors' choices. The payoff structure considering the network thus become:[8]

$$\begin{aligned} \pi_i(x_i, \vec{x}_j) &= \sum_{j \in d_i} \left(V - f(x_i - x_j)^2 - (1-f)(x_i - t_i)^2 \right) \\ &= k_i \left(V - (1-f)(x_i - t_i)^2 \right) - f \sum_{j \in d_i} (x_i - x_j)^2 \quad . \end{aligned} \tag{1.3}$$

where the unique best response for agent i is

$$x_i^*(\vec{x}_j) = f\mu_i + (1-f)t_i \quad , \tag{1.4}$$

[6] This is a simplification. However introducing an extra layer of heterogeneity in this dimension would only complicate the nature of the results, without delivering interesting insights.

[7] See Bolletta (2015) for the study of a similar model under undirected network. There it is shown that it is still feasible to perform the analysis, although some extra conditions are required in terms of rationality to refine multiplicity of equilibria in the network formation stage. In that paper farsightedness is used as a refinement. We preferred here to stand to a more realistic assumption on individual's rationality, which in addition allows us to solve the model through the concept of Nash equilibrium.

[8] Note that although we focus on the following payoff structure, what really matters for the results that follow is the best reply scheme. Indeed one could redirect the analysis that follows to a payoff structure consistent with other parametric forms already used in the literature, such as Calvó-Armengol et al. (2009) among others.

where we have called $\mu_i \equiv \frac{\sum_{j \in d_i} x_j}{k_i}$. This is equation 1.1, already described in the introduction. At this stage we analyze one period only, and in a later section we discuss dynamics of the model.

The timing of the on period game is as follows.

Definition 1. *TIMING:*

Time 0: *Opinions $\vec{t}$ are exogenously assigned to agents;*

Time 1: *agents form simultaneously and independently the directed network, which becomes common knowledge;*

Time 2: *opinions are updated into x, according to the network and to best replies.*

From the timing we introduce the equilibrium concept. Formally we see this as a sequential game, that we can therefore solve by backward induction. In particular we focus now in the solution of the system of best replies, when the network is given, and therefore μ_i is uniquely defined for all i. Then in the next section we move to the previous stage letting agents form the network, to finally move to the very initial stage on the distribution of opinions to characterize the possible equilibria.

To proceed, let us introduce some more notation to rewrite the system of best replies in matrix form. To do so we call:

- F the diagonal matrix of all flexibilities, so that $F \equiv \begin{pmatrix} f_1 & & 0 \\ & \ddots & \\ 0 & & f_n \end{pmatrix}$;
- $\vec{t}$ the vector of all types;
- and D the adjusted adjacency matrix such that $D_{ij} \equiv \begin{cases} \frac{1}{k_i} & \text{if } j \in d_i, \\ 0 & \text{otherwise.} \end{cases}$

Note that we let f be homogeneous, although the model could account for heterogeneity in this dimension. Moreover we underline once again that the entries into the matrix D have to be determined in equilibrium. Then a compact way to write (1.4) is

$$(I - FD)\vec{x} = (I - F)\vec{t} \ . \tag{1.5}$$

Lemma 2. *Equation (1.5) has a unique solution $\vec{x} \in [0,1]^n$.*

Proof. See Appendix. □

This simple result states that for a given network there is a unique Nash equilibrium of the game. This is crucial because we can now move forward and go study the network formation. Before doing that we focus on the payoff structure and derive a formula for the *payoff in equilibrium*, which dramatically helps us in the analysis of agents' strategies. Thus let us move back to the Nash equilibrium. From (1.3) and (1.4), the payoff in equilibrium is[9]

$$\pi_i \quad = k_i \left(V - f(1-f)(\mu_i - t_i)^2 - f\sigma_i^2 \right) \tag{1.6}$$

where we have called $\sigma_i^2 \equiv \frac{\sum_{j \in d_i}(x_j - \mu_i)^2}{k_i}$ the variance of the actions of i's neighbors. Given that the payoff structure is quadratic, it is not surprising that second moments of the distribution of behaviors appear in the analysis. Nevertheless, this is a result that is not highlighted from previous works, although it is particularly meaningful. Interestingly, we see how fully flexible agents ($f = 1$) have preferences only for homogeneous groups, and they would not care about which opinion the group exhibits. Therefore they will form the group with agents that share the most similar opinions. For agents that instead are not prone to change their opinion ($f = 0$), others' opinion are completely irrelevant, and therefore connections are formed at cost 0. For intermediate values of flexibility ($f = 1/2$), agents both care about having homogeneous and similar opinion groups. Next Remark formally states some comparative statics on f on the payoffs.

Remark 3. *The payoff of an agent in equilibrium depends:*

- *quadratically on μ_i, with a maximum when $\mu_i = t_i$ – this effect is the most detrimental when $f = \frac{1}{2}$;*
- *linearly on σ_i^2 – this effect is the most detrimental when $f \to 1$.*

So, the payoff seems always maximum when $f = 0$, so that $x_i = t_i$ (but it could be a problem to find neighbors when the network is endogenous).
The first order effect (i.e. fixing others' best responses) of increasing flexibility up to $\frac{1}{2}$ decreases welfare, but when $f > \frac{1}{2}$ an larger f could increase welfare if the actions of neighbors have low variance.

[9] See Appendix for the complete derivation

1.3.2 Endogenous network

The concept of equilibrium here is sub-game perfect Nash equilibrium. Since the network is directed the solution is fully non-cooperative, and it is worked out in two stages, by backward induction. In fact the agent will first choose their neighbors, and then will update their behaviors according to equation 1.4. In the previous section we established uniqueness of equilibrium for a given network and we can therefore focus now on the network formation stage.

Before moving to the result, we propose here a definition that allows us to partially characterize the equilibrium.

Definition 4. *An equilibrium is **ordered** if each agent in a network g has an interval of neighbors $\{i - a_i, i + b_i\}$ such that $a_{i+1} \leq a_i + 1$ and $b_i \geq b_{i+1} + 1$.*

The above definition can be described as follows. We call an equilibrium to be ordered if agents are matched only with their closest neighbors. Clearly every agent will have different bounds on their neighborhood, and we could characterize those only departing from a given distribution of opinions. The following result formally states that this configuration is the unique that can arise in equilibrium.

Proposition 5. *An ordered equilibrium exists and every equilibrium is ordered.*

Proof. See Appendix. □

To get an intuition on this result, let us recall equation 1.6. In particular we could broadly interpret the payoff deriving from a set of links as follows. First agents want that the average behavior in their neighborhood is close enough to their initial opinion. That is agents are *homophilous.* Therefore agents with similar opinions will tend to match together. This already pushes strongly towards uniqueness of the equilibrium. In addition to that we should consider the second term of equation 1.6, which shows how agents have preference over the diversity of opinions inside their neighborhood. Therefore we can say that agents will tend to match with their most similar agents, and the reference groups are well defined since the variance term ensures sharp bounds on it.

Given the previous result we can now determine the conditions on the initial distribution of opinions such that the network will be disconnected. In particular we will have that the spread of opinions is greater than the initial distribution of opinions if the network is disconnected, while lower if instead the network is connected.

The model allows to define a triple $\langle V, f, T\rangle$ which will map into a network configuration and a distribution of opinions $\langle G, X\rangle$. Thus it is a matter of interest to understand under which condition the distribution of ex-post opinions is more polarized than the distribution of initial opinions. Simply defining a Gini type of measure, we can compare the vector $\mathbf{x}$ with the vector $\mathbf{t}$. As it turns out, the weighted adjacency matrix is all matters here. Such a result is therefore comparable with all the literature on DeGroot processes. We need to find condition on aperiodicity, irreducibility and other Markov Chain results for the adjacency matrix.

To test for the polarization into the population of interest, we should first define a Gini type coefficient for our specific context. To do so, let us introduce a vector

$$\bar{\xi}' = (-n+1, -n+3, ..., n-1, n+1)$$

Thus, if we calculate $\bar{\xi}'\bar{x}$ with $\bar{x}$ being the vector of actions in equilibrium, we have a measure of dispersion of actions. Clearly this will not be between 0 and 1 as the typical Gini index, since we would need to weight it, but we are rather interested in the difference in dispersion between equilibrium behaviors $\bar{x}$ and initial opinions $\bar{t}$. We want then to analyse the sign of our measure of polarization $P(g,t) = \bar{\xi}'(\bar{x} - \bar{t})$. If positive we would have polarization, if negative we would have centralization, and if 0 then we would have that the endogenous network would not matter at all. Interestingly, if we use equation 1.5, we could simplify this measure, as shown by the following simple algebra.

$$\begin{aligned} P(g,t) &= \bar{\xi}'(\bar{x} - \bar{t}) \\ &= \bar{\xi}'[I - \frac{f}{1-f}(I - fD)]\bar{x} \\ &= \frac{f}{1-f}\bar{\xi}'[D - I]\bar{x} \end{aligned} \tag{1.7}$$

From this we see clearly how the sign depends only on D, while it is independent of f. In particular the equation above is going to be positive if D is disconnected, negative otherwise. This is to put emphasis once more on the relevance of the network formation stage in the process analyzed here. Provided that, we move now to the main result of the paper, which provide sufficient conditions on the distribution of opinions such that the network is going to be disconnected in equilibrium.

Proposition 6. *If there exist two agents i, j such that $j = i + 1$ and $|t_i - t_{i+1}| > \xi(f, V)$, all equilibria are disconnected.*
Such threshold is $\xi = \left(\frac{(1+f)^2}{f(1-f)}V\right)^{\frac{1}{2}}$, and is convex in f, with a minimum in $\frac{1}{3}$.
Equilibria exhibit a number of components $|C| \geq \{i, j|\xi < |t_i - t_{i+1}|\} + 1$

Proof. See Appendix. □

From the above result we learn that first of all there exists a threshold on distance of types such that, for all rules of order in the network formation stage, the network will be disconnected. This is therefore a sufficient condition.

Importantly, the result holds for any distribution but the uniform. To see that, it is enough to point out that we must check among all couple of agents, and determine the couple characterized with the maximum distance. Then we can check that there always exist a pair of V, f such that that distribution may exhibit disconnectedness. Finally note that every distribution that is not-uniform, grants the existence of a couple of agents characterized by maximal distance in types.

Interestingly the threshold is convex in f, with a minimum for $f = 1/3$. The intuition behind this result is that agents have to weight enough the variance term. However, there is a non monotonic effect for larger f since agents will also be more influenced by neighbors, so overall variance would be lower. Nevertheless this implies also that in the first place agents prefer to have homogeneous neighborhood, and that is why in equilibrium they are likely to take place locally, with segregation arising.

In the next example we show simulations of the model for several distributions, showing how the uniform would be strongly connected, while other distributions, namely a bimodal and a normal, favor the arousal of more groups not inter-connected.

1.4 Dynamic setting

In this section we extend the result shown before to a dynamic setting. In particular we let that at any point in time an agent's opinion correspond to her ex-post opinion, determined in equilibrium in the previous period. Formally we have then $t_{i,s} = x_{i,s-1}$, where $s \in \{0, 1, .., S\}$ denotes discrete time and with $t_{i,0} = t_i$. Clearly we want the process to be myopic. In fact agents at every step s will only maximize utility at that given point in time. Intuitively, the previous analysis focus on what happens within a time step, and now we consider multiple time steps.

Formally the timing is now given by:

- Agents get types $s = 0$
- Then form the network and update them $s = 1$
- At every following stage they inherit the type from previous period $t_{i,s} = x_{i,s-1}$
- Keep on updating the network and opinions

In this way we are able to initialize the model with a distribution of times and study both the long-run behavior of the society and the step-by-step evolution of the process. In particular we are able also to see deeper how the result obtained before relate to the well known results found in DeMarzo et al. (2003) and Golub and Jackson (2010). This papers show how if the network is strongly connected and weighted adjacency matrix is primitive and aperiodic (as it turns out the latter implies the former), consensus is always reached.

Our model could be in fact compared to the model by Krause (2000) and Hegselmann and Krause (2002), where he shows that if agents give weight only to those who have similar opinions, convergence of opinions will be ensured only within each component, even if the network is strongly connected. Somehow in these models connections do not matter, as it matter instead the agents to which it is given weight in the process of opinions updating. The rule to determine the closest neighbors to which pay attention is imposed exogenously, while again our model generate that in equilibrium through the endogenous formation of the network.

The dynamic process is characterized by the one-step transition already analyzed, and after initializing the process at $x_{i0} = t_i$ the process is described by the following.

$$x_i(S) = \sum_{s=1}^{S} f(1-f)^{S-s}\mu_i(s) + (1-f)^S t_i$$

where both $x_i(s)$ and $\mu_i(s)$ are determined at each point in time according to the above process.

Remark 7. *The dynamic model exhibit the following properties:*

- *For each $k \in C$, $x_k(s)$ converges to a x_k^*.*
- *For each h, k, either $x_h^* = x_k^*$ or $|x_h^* - x_k^*| \geq \xi$.*
- *For each $k \in C$, $g_k(s)$ is the complete graph.*

The result follows from the analysis in Hegselmann and Krause (2002), with the only difference brought in by the endogenous network (the third point), and therefore we omit a formal proof.

To see convergence to the complete network it is easy to observe that along the path behaviors tend to condense, and therefore locally there will be less dispersion. This feeds into more link formation, which again brings additional inertia to the whole process. Nevertheless, we have that if the network is disconnected at the first time step it will remain disconnected, and viceversa. Consensus will take place only partially within every component that arouse from the network.

Accommodating our model to a dynamic version we see that all that matters for convergence is the formation of the network in the first period. Then one or more components arise accordingly, and over time we observe a contraction of opinions toward a unique one, in each component that has arouse. At the same time, the network will become more and more connected over time, converging to the complete one.

The endogenous network shapes the adjacency matrix into blocks, and then over time there is an increase in density of positive entries because the actions are closer and closer. As a result we get a reinforcement over the process of convergence. We can surely say that an increase in V and in f would push both over a faster convergence. This is intuitive, but considering the flexibility, we know also that it makes easier to break up the network.

1.4.1 Simulations

In this section we provide some numerical results for the dynamic version of the model. We provide results for two different distributions of types, keeping all the other parameters fixed. In particular we fix the following choice of parameters. Given the parameter choice

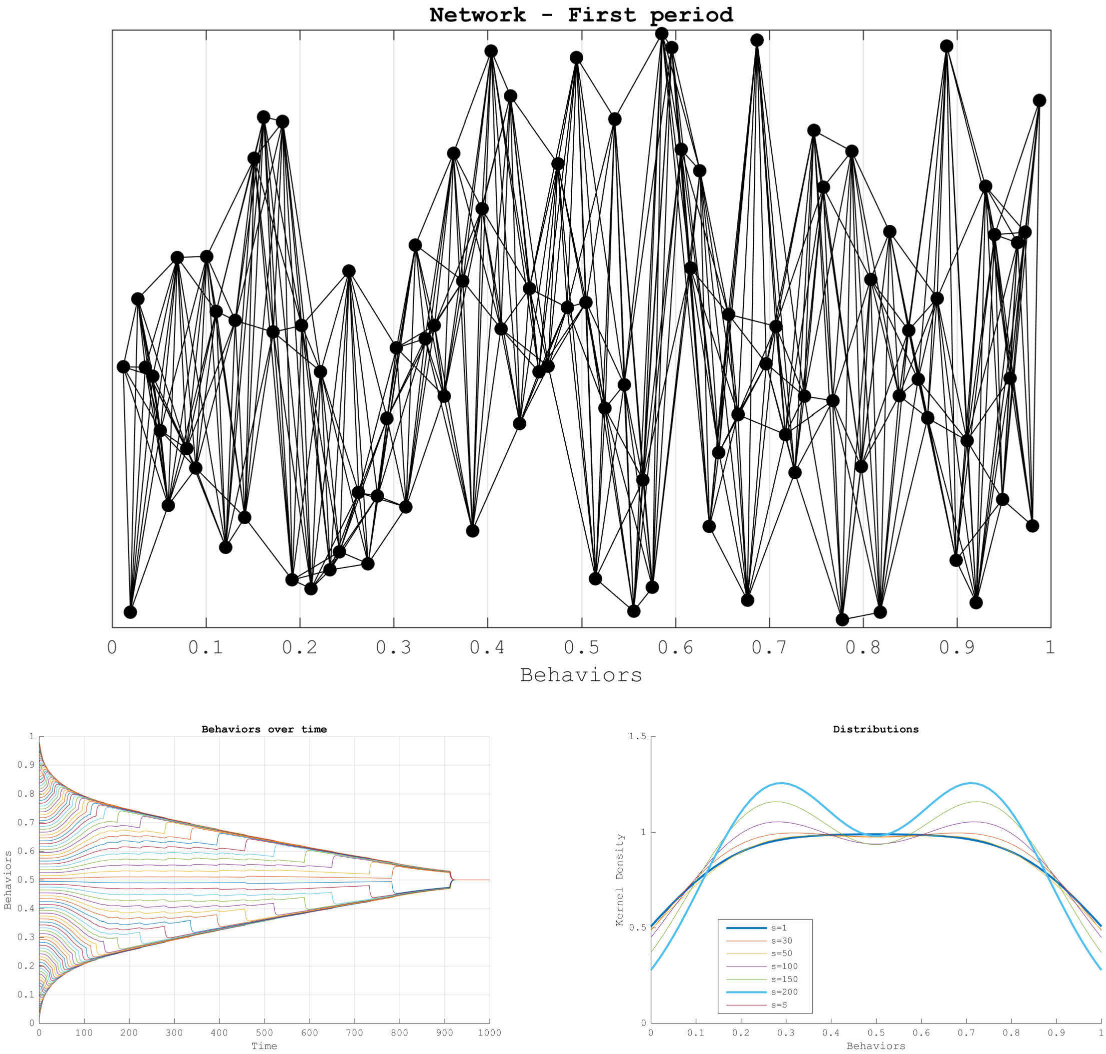
Parameter	**Value**
N	100
f	0.3
V	0.0007

the threshold is identified with $\xi = 0.078$.

Then we are interested in one-shot and long-run behavior of different distributions. We first analyze as a benchmark a uniform distribution. As we learn from the above result such network is strongly connected and it will eventually converge to a consensus over time. The second distribution instead uses a distribution generated from a truncated bimodal, defined as the mixture of a truncated normal $g_1(t) \sim \mathcal{N}[0, 0.4]$ and $g_1(t) \sim \mathcal{N}[0.48, 1]$. Thus we forced the data to exhibit at least two components.

Let us comment first the figure 1.1. The top graph show the equilibrium network selected by the solution algorithm, and we see that it is connected. So the process must converge in the long run. The most interesting remark is observing the transition toward the convergence point, evidenced by the two figures below. In the bottom left figure we represent the evolution of behaviors over time. We observe mainly two things: *i)* the process takes a remarkable amount of time to get to consensus, and *ii)* there is a

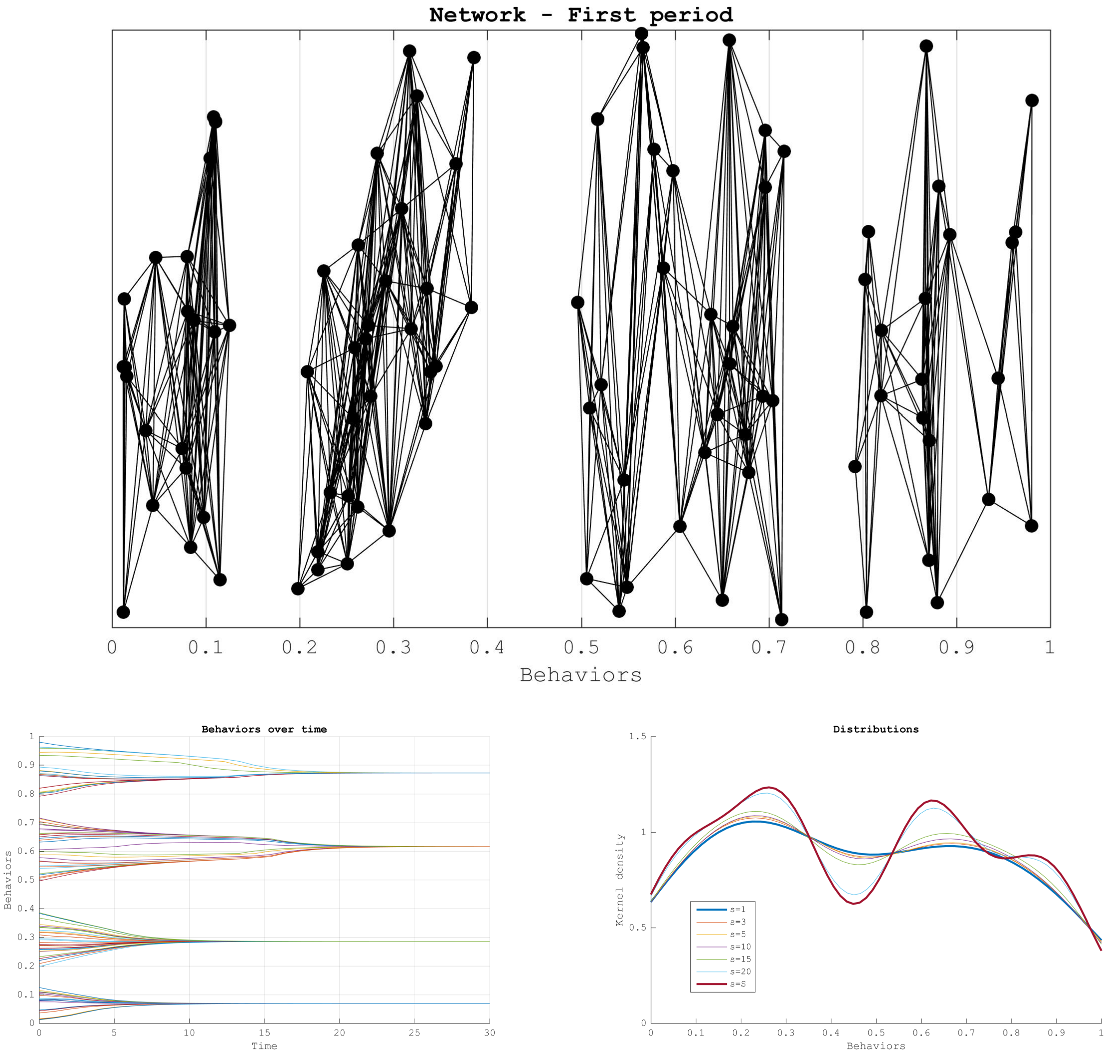
Figure 1.1: Uniform distribution



condensation of opinions on the tails of the distribution. More and more agents will get toward extreme behaviors, although the extremes themselves get closer over time. We call this feature transitional polarization. This is documented by the figure in the bottom right, where indeed we notice that from the initial uniform distribution, where by definition polarization is absent, there is a consistent polarization process over time. In the very long run it will disappear and eventually the system converges toward a unique behavior. However, over the path it is remarkable how polarization may arise. We believe this result is novel and particularly interesting for several applications.

Figure 1.2 differently show a distribution that we forced to exhibit at least two components. Random sampling of data ended out exhibiting 4 of them. What we observe is how partial convergence takes place within every component, and how also here the process in the end will exhibit a maximal amount of polarization given the averages of

Figure 1.2: Random sampling from bimodal distribution



each subgroup. What is remarkable to notice is also that here the process is consistently faster than in the previous case.

1.5 Discussion

In this paper we showed how through the endogenous network tend to be disconnected. In particular this happens when forming connections is not too profitable, and agents are particularly sensible to social influence from their neighborhoods. This conditions lead opinions in the long run to diverge, and thus consensus is never reached.

In several real world examples the choice from the agents of their neighbors is non-negligible. From this results we learn therefore that considering a network formation process has dramatic results on several dimensions, namely the network configuration, equilibrium behaviors, long run opinions and speed of convergence.

We wanted to contribute in this sense to the existing literature on opinion formation processes, because this could really open new research questions. In particular it comes natural the question regarding possible nuisances that may alter the distribution of opinions at any given point in time. This could explain cyclic behavior of political parties, strategic provision of information from media and evolution of social norms. These are relevant questions well explored in the literature, but we believe that this work can consistently contribute having a richer understanding of them.

Chapter 2

Peer effects in school: a model[1]

In this paper I develop a model of network formation that exhibits peer effects. Students form friendships, and their academic outcomes will depend on the acquaintances they created within the class. I allow for the presence of role models, to identify the students with more popular characteristics. Therefore I model the problem of a school, who gets a number of freshmen and must decide to assign students to classrooms. The school's objective is to maximize student's outcomes. Numerical results, consistently with experimental evidence, show that if there are positive role models the school must create uniform classrooms, isolating negative role models.

[1]I would like to thank in particular my supervisor Giulio Zanella, Paolo Pin and Andrea Galeotti for their insightful comments. Moreover a special thanks goes to my Ph.D. colleagues in University of Bologna for the great time spent together, the attendants of the Internal seminar series in University of Bologna, as all the participants to the NEXT Workshop in Budapest (2014), the Winter Workshop on Networks in Economics and Finance in Louvain-là-Neuve (2014) and the Networks Discussion Group in Stanford University for their constructive questions that helped me improving this work. All errors and mistakes remain my own.

2.1 Introduction

In a society men are supposed to interact together, and to internalize the needing and the characteristics of surrounding agents. Otherwise we would be only islands. On the other hand we choose to be surrounded by specific individuals, and even though casualties play certainly a role, most of the social relation are part of the decision process of the individual. The main source of complexity, and at the same time its own beauty, is due by the fact that we are all endowed with our own identities.

Among the first to introduce the role of identities into game theoretical settings there is the paper by Akerlof and Kranton (2000). The authors have emphasized with several examples how individual heterogeneity in the feelings with the self could build interesting extensions to the already existing models of strategic choice. In this paper I want to explore the role of this dimension both in the shape of agent's behaviors and in the choice of their connections.

The main findings are that agents in equilibrium will exhibit homophilous behavior, and this depends on the best replies structure. The intuition behind is that agents enjoy behaving according to their identity, therefore choosing an action in line with that, but at the same time they have to mediate their actions in order to be a member of a reference group. To accomplish these findings I derive a characterization result that holds for all pairwise equilibria in this context, and reduce the set of feasible networks to those that show patterns described in Watts and Strogatz (1998) as *regular networks*. However we know that pairwise stability is a subtle equilibrium concept, as it usually identifies a large number of networks, and cannot avoid the existence of cycles, as proved in Jackson and Watts (2002). Especially to avoid the latter I refine multiplicity of equilibria applying the concept of *farsightedness*, applying the definitions of Herings et al. (2009). In this context this refinement tool is particularly effective in avoiding cycles, even though in principle it is not a refinement of pairwise stability. Nevertheless, *farsighted pairwise stable networks* turn out to be unique.

To achieve these results I developed a network formation model with simultaneous determination of equilibrium behaviors. Actions are defined through a continuous variable, that is determined as a function of other players' behaviors and agent's own type. In particular the agents that enter in the utility function are those with which a link is formed, which in turn is the initial endogenous choice. The model is solved as in two stages, given that the subgame delivers a unique Nash prediction and therefore agents can choose the preferred network updating their behavior according to that. Again the main forces are that agents want to stick to their identity, but forming some social relations, which is profitable, they have to adapt to other agents' actions. These ingredients lead agents to form connections with more similar agents.

Just to make an intuitive example, imagine a group of friends that really love rock music. When they meet they will talk a lot about that and how good it is. If one of these individuals hates rock music, she would feel bad both if she ends up talking about rock music, but also if she would say that rock music sucks. This would be why, in the first place, it is extremely unlikely that the individual who hates rock would hang out with rock music enthusiasts. These are all mechanics explained by the model.

Then the model is studied under particular frameworks, so to suggest some applications. First I study a case in which there are two cultures in a society, and there is a majority and a minority. I focus on conditions under which the two societies would ever form a connection between themselves. As a result the willingness of forming connections depends on the relative shares of the two groups in the whole population. If the minority is "small enough", then we may have that the majority is willing to form connections with the minority, because it is profitable for them, but the minority refuses such connection. Therefore by pairwise stability, the two cultures will not be connected, and no transmission of culture takes place. I believe this result is novel, and even though it sounds counterintuitive there is lot of anecdotal evidence of minorities refusing to absorb the social norms of their hosting countries. Some suggesting evidence is provided, and it derives the policy implication that in some cases it would be preferable to open boundaries, rather than close them, because it could foster integration of communities. However, this is a result that calls for more empirical evidence.

As a second application, I analyze a context of peer effects in the classroom. There are lot of factors behind this particular effects, but I want to focus on the mechanics of the proposed model. Nevertheless, to improve interpretative power, I differentiate this exercise letting the benefit from the interaction be a monotonically increasing function of types, that reflects the idea for which better students receive more friendship requests, as agents may strategically want to link with them to exploit their influence in succeeding in school. The results in this context are quite rich of implications. If the distribution of student abilities is smooth (for instance as in a randomized classroom) connections are likely to arise, and low ability students would benefit from indirect transmission of positive peer effects, while best achievement students would not suffer from negative influence. If instead the students' abilities are not evenly distributed, segregation is likely to arise, and low ability students would not benefit from peer interactions. The results in this section are discussed benefiting from existing experimental literature, which proves also real world relevance of the model developed.

Being a context of peer effects I also show that the best replies system captures the main ingredients of the linear-in-means model by Manski (1993). Moreover from Calvó-Armengol et al. (2009) we know that peer effects may strongly depend on Bonacich centrality, and I show that this is also a feature of the model, with the main difference that here agents choose how to be central in the network, because the link formation is fully strategic. To favor the comparison with existing literature I also discuss how the results hold under the functional form used by Calvó-Armengol et al. (2009),[2] with the due modifications to allow identities enter the model. Best replies are the same exploited in the main specification of the model, and in fact they are the main characterizing ingredient of this framework.

The rest of the paper is organised as follows. Section 2.2 review the existing related literature. Then in Section 2.3 the model is extensively discussed, and all results are shown. Section 2.4 shows the three exercises mentioned above. Finally Section 2.5 discusses the main implications of the model as well as possible extensions. All proofs are in the Appendix A.

2.2 Literature Review

This paper aims to contribute to the literature on strategic network formation. In the first place the paper by Jackson and Wolinsky (1996) proposed a framework to work with, and introduced the notion of pairwise stability. From that paper I took inspiration from the so called *connections model*, which elicits strategic behaviors through an extremely reduced form of benefit and costs deriving from interactions. Here I go beyond this reduced form and I let the costs deriving from interactions be determined by an underlying coordination game. The structure of the problem characterizes a network game. The usual practice in the literature was to study games in this category keeping the network exogenous. For instance Ballester et al. (2006) studies a game of strategic complements, while Bramoullé and Kranton (2007) and Bramoullé et al. (2014) focus on strategic substitutes. Making the network endogenous is instead relatively new, and to the best of my knowledge few works attempt to do so. Among these there is the paper by Galeotti and Goyal (2010) which is the first contribute into the literature in this direction. More recently there is Kinateder and Merlino (Kinateder and Merlino). Both these works analyze public good games, and therefore games with strategic substitutes.

[2]See Appendix B for the discussion of this aspect.

Among the main implications of the model is that homophilous behavior arises in equilibrium. This is a topic that was largely investigated in the literature on networks, as it is one of the features that characterizes complex networks, and that do not arise in random networks. Among the several works there is the paper by Currarini et al. (2009) which evidences out the role both of preferences in link formation process, and of randomness. Agents have preferences in linking with same-type agents, but also chances of meetings (i.e. relative availability of same-type agents into a larger population) matter. Using a really different approach, Golub and Jackson (2012) show how homophily can influence best-response dynamics and learning processes. For higher importance of homophily into a society, it is less likely that such society will reach consensus. More recently Boucher (2015) developed a network formation model, where homophily is a key component and enters the analysis through a distance function on agents' characteristics. The paper proposed here extends this framework adding equilibrium behaviors on top of a linking formation framework, and homophily arises endogenously.

Regarding the methodology this paper provides an application of the concept of far-sightedness in networks, introduced by the works by Page Jr et al. (2005) and Herings et al. (2009). The latter is the paper I strongly rely on in terms of definitions. The application of such concept is, to the best of my knowledge, rare. One of the reasons why it is so, I believe on the strong rationality that imposes to agents. In fact, through this concept, agents must be able not only to take into account others' strategies, but also to internalize reactions to their actual choices. In fact agents can strategically undertake a non optimal action if this leads, through optimal adjustments from other players, to a preferred outcome. Especially when the network is large this implication may not be credible. Nevertheless, one could obtain comparable results through an *ad hoc* dynamic process of best response learning, which is much more frequent to observe.

I finally mention some related literature on the applications I propose in this paper. The latest works on peer effects in the classroom exploit composition of classrooms. Keeping randomized classrooms as benchmark, the treatment consists of specific compositions of students' body that are expected to enhance transmission of peer effects. Nevertheless, no theory is behind those treatments, and moreover they have not delivered the expected results in some cases. For instance the work by Carrell et al. (2013) identified "bimodal distribution" of students as those more likely to foster peer effects, targeted on low achievement students. The results showed that students reacted endogenously to this treatment in terms of group composition, and the two group of students (low achievement and high achievement) remained segregated. In this paper I propose a theoretical framework that could be exploited in that direction, and therefore could guide further experiments having clearcut predictions as a benchmark. Other examples in this literature are Li et al. (2014) which to favor interaction among high achievement and low achievement students through monetary transfers. This is interpretable in the framework of this paper as an increase on the benefits deriving from the interactions.

2.3 The Model

2.3.1 Notation

Let $N = \{1, 2, ..., n\}$ be a finite set of agents with $n \geq 3$. Every player is endowed with a type θ drawn from a distribution with $\theta \in [\underline{\theta}, \bar{\theta}]$. This is to introduce a dimension of heterogeneity of players. Potentially it could represent a vector of idiosyncratic characteristics that makes every agent different from another, but here we let θ be mono-dimensional. The network relationships are reciprocal and the network is represented through a non-directed graph. A network g is a list of which pairs of individuals are linked to each other. We write $ij \in g$ to indicate that i and j are linked under network g. The set of all possible network or graphs is denoted by $\mathbb{G}$. The network obtained adding link ij to the existing network g is denoted with $g+ij$, and the resulting network deleting link ij is denoted with $g-ij$. For any network g let $N(g) = \{i|$ there is j such that $ij \in g\}$ be the set of players who have at least one link in the network g, and for every i let $N_i(g) = \{j|ij \in g \; \forall \; j \neq i\}$ be the set of links for an agent i, i.e. her neighbourhood. $|N(i)|$ is the cardinality of the set of peers of agent i, i.e. their degree[3], denoted throughout the paper with $d_i(g)$. A path in a network $g \in \mathbb{G}$ between i and j is a sequence of players $i_1, ..., i_k$ such that $i_k i_{k+1} \in g$ for each $k \in \{1, ..., K-1\}$ with $i_1 = 1$ and $i_k = j$.

2.3.2 Payoffs and Preliminary Results

The choice of agents is an action defined x on a continuous space $\mathcal{X}$, that we can interpret here as an ex-post action, after social interactions took place. Under this framing the identities would be the ex-ante behavior.[4] As standard in a local interaction framework there is a mapping of actions and types into utilities. Formally we have that $(x_i, x_j, \theta_i) \mapsto U_i(x_i, x_j, \theta_i)$. A parametric representation of such an environment is given in Bisin et al.

[3]Since the network is undirected there is no need to distinguish between in-degree and out-degree, because in such context they coincide.

[4]This is not an assumption, but as later discussed equilibrium behavior in isolation is that an agent responds only to her identity, and no social influence takes place.

(2006), through the following utility function

$$U_i(x_i, x_j, \theta_i) = -\alpha(x_i - x_j)^2 - (1-\alpha)(x_i - \theta_i)^2 \tag{2.1}$$

where i, j indicate agents with $i \neq j$, and α is a weighting parameter exogenously given to all players. In order to avoid multiplicity of equilibria at this stage we need that $\alpha \in [0,1)$, because otherwise behaviours may be driven only by social interactions.[5]

This function, though, allows only to consider bilateral relations among players, while here I am interested in multilateral connections and their strategic formation. Formally I want that also the network is in the utility function, and so $(x_i, x_j, \theta_i, g) \mapsto U_i((x_i, x_j, \theta_i, g))$. Therefore I will change slightly this utility function in a way that it can represent (i) the benefit agents can get from direct and indirect links, a la Jackson and Wolinsky (1996) and denote it with δ, (ii) a sum of cost agents pay from deviating from peer's behaviour, allowing then for multiple links (cost of conformism), and (iii) a weight players give to their peers, which in this paper we assume to be equal among all agents and to $\frac{1}{d_i(g)}$. This choice can be interpreted as the intensity of the interactions. If an agent has relatively few links, she would be more influenced from her links than any other agent having more links.

The preferences of an agent $i \in N$ can therefore be described through the following.

$$U_i(x_i, \{x_j\}_{j \in N(i)}, \theta_i) = \sum_{j \in N_i} \delta - \alpha \frac{1}{d_i(g)} \sum_{j \in N_i} (x_i - x_j)^2 - (1-\alpha)(x_i - \theta_i)^2 \tag{2.2}$$

where $\alpha \in [0,1)$, and it is homogeneous among agents. $\delta \geq 0$ is equal for all agents.

The chosen utility function is an isomorphic transformation of the equation 2.1, and therefore the introduction of the benefit is irrelevant to the choice of actions, allowing us to evaluate the best replies easily, because they are linear. Although, I am not able in this way to determine simultaneously also the reference groups, neither in their size nor composition. To do that we need to workout the full equilibrium of this game in two stages. Note that for any given reference group, the second and third terms of the right hand side of equation 2.2 are determined and fixed, and therefore can be considered as the cost of keeping links, that as standard agents would pay only for every direct link they are willing to form. In this sense this paper extends the standard framework of Jackson and Wolinsky (1996), as it allows for optimal behaviours to be simultaneously determined,

[5]See Lemma 8 for a formal statement of this result.

and studied in the equilibrium.

Best replies here are simply given by the following.

$$x_i^\star = \alpha\{\bar{x}_j\}_{j\in N_i(g)} + (1-\alpha)\theta_i \tag{2.3}$$

The existence and uniqueness of the equilibrium in the coordination game represented above is guaranteed by linearity of best replies. Next Lemma formalizes the result.

Lemma 8. *For all $0 \leq \alpha < 1$ there exist a unique Nash equilibrium of the coordination game with $x_i^* \in [\underline{\theta}, \bar{\theta}]$ for all $i \in N$.*

Proof. See Appendix. □

This result basically states, that for every possible exogenous network, the vector of equilibrium behaviors is unique. Provided that we can now move to the endogenous network.

2.3.3 Endogenous Network

Allowing the network to be endogenous, agents would then choose the set of links, and then update their behaviors according to the findings of equation 2.3, which are unique thanks to Lemma 8. We have that $\mathcal{X}_i = [\underline{\theta}, \bar{\theta}]$, and $g_i = \{g_{i1}, g_{i2}, ..., g_{iN}\} = \mathcal{G}_i$. Therefore the individual strategy is represented by a couple $s_i = (x_i, g_i) \in \mathcal{S}_i = [\underline{\theta}, \bar{\theta}] \times \{0,1\}^N$. However the network is undirected, therefore we cannot rely on perfect subgame Nash equilibrium, but I will look at pairwise Nash stability as defined in Calvó-Armengol and İlkılıç (2009). This definition extend the standard pairwise stability framework introduced by Jackson and Wolinsky (1996) because it allows the agents to sever multiple links at once. In fact, the rigidity implied by pairwise stability does not allow to refine enough, and basically too many networks are selected as possible equilibria, not allowing any characterizing pattern to emerge.

To approach the result on stability I first focus on characterizing the equilibrium, which will also be useful in the further analysis. Studying equilibrium behaviours we see that, first of all, every agent has a strict preference in interacting with most similar agents. Therefore, agent 1 would never make a link with agent n if there is not already an existing link with agent $n-1$, because in such a case, 1 would strictly prefer to link with $n-1$. These features are stated formally in the following Proposition.

Proposition 9. *For any $\delta, \Theta, \alpha \in [0,1)$, every pairwise Nash equilibrium is characterized by the following linking scheme:*

- *if* $g_{i,i+k} = 1$ $(g_{i,i-k} = 1)$*, then* $g_{i,i+(k-1)} = 1$ $(g_{i,i-(k+1)} = 1)$*,*
- *if* $g_{i,i+k} = 1$ $(g_{i,i-k} = 1)$*, then every* $j \in \{i+1, ..., i+(k-1)\}$ $(j \in \{i-1, ..., i-(k-1)\})$ *are such that* $g_{j,i+k} = 1$ $(g_{j,i-k} = 1)$*,*

$\forall i \in N$ *and* $\forall k$*.*

Proof. See Appendix. □

From this result we learn that the network, if driven by the underlying game described above, will exhibit a particular structure. In fact this is the network obtained by Watts and Strogatz (1998) when having the "regular network".

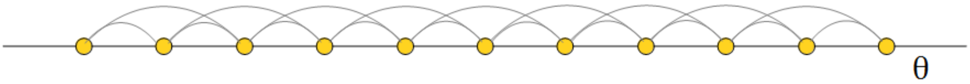


Figure 2.1: The regularity of the network.

Proposition 9 provides a strong characterization and, moreover, because reduces dramatically the set of possible networks we could study. Intuitively it shows that there is a sort of sequentiality in choosing peers, depending on social distance. This mechanism is also known as inbreeding homophily.

The parameter α turn out to be particularly relevant, because it is among the main driving forces. Therefore next result states that it is more interesting to focus on middle values of the parameter, since at the boundaries the predictions of the model ar rather stark.

Remark 10. *When* $\alpha \to 1$ *or* $\alpha \to 0$ *the complete network is strongly efficient for any* δ *and for every distribution of* θ*.*
In such cases the complete network is also uniquely pairwise Nash stable.

Proof. See Appendix. □

Previous Remark states that analysis with extreme values of the weighting parameter α becomes trivial. For such extreme values it is like the cost of connections are null, and therefore adding a connection is always beneficial.[6]

[6]Note that if we would have added a fixed cost into the utility function we would have replicated exactly the results in Jackson and Wolinsky (1996).

Thanks to the results presented above we obtained a characterization of the set of possible pairwise stable equilibria. Nevertheless, we cannot state anything about uniqueness, yet. In particular we know that a pairwise stable network can fail to exist when a cycle arise. Therefore studying the static environment we can fail to have a pairwise stable equilibrium. However I want to achieve uniqueness as it would allow to perform comparative statics exercises and to study specific contexts. To do so, I will move towards a definition of stability that uses dynamic reasoning, letting the system to be static. In simple words we could say that agents performs the dynamics in their mind, allowing then to undertake links choices out of the pairwise stable scenario, if that leads to a network that leads them better off. This is what we mean by *farsighted* agents.[7] Note that this is not a refinement of pairwise stability, as it can determine an equilibrium in presence of a cycle, where pairwise stability fails to exist. The following part of this section is devoted to the derivation of the result on equilibria, which ensures that there are no cycles.

I introduce here some definitions needed for the result. First of all I report the definition of farsighted improving paths,[8] found in Herings et al. (2009)[9]

Definition 11. *A farsighted improving path $F(g)$ from a network g to a network $g' \neq g$ is a finite sequence of graphs $g_1, ..., g_K$ with $g_1 = g$ and $g_K = g'$ such that for any $k \in \{1, ..., K-1\}$ either:*

- $g_{k+1} = g_k - ij$ *for some* ij *such that* $U_i(g_K, \theta_i) > U_i(g_k, \theta_i)$ *or* $U_j(g_K, \theta_i) > U_j(g_k, \theta_i)$ *or*
- $g_{k+1} = g_k + ij$ *for some* ij *such that* $U_i(g_K, \theta_i) > U_i(g_k, \theta_i)$ *and* $U_j(g_K, \theta_i) > U_j(g_k, \theta_i)$.

Now we introduce the concept of pairwise farsightedly stable set. The formal definition is found in Herings et al. (2009)[10], and I report it to accomodate the slightly different notation.

[7]Other equilibrium concepts have been explored. Among others, Strong Stability would have been a strong equilibrium condition. However strong stability relies on the existence of a potential function. As shown in Monderer and Shapley (1996) such function exists if and only if $\frac{\partial^2 U_i}{\partial x_i \partial x_j} = \frac{\partial^2 U_j}{\partial x_j \partial x_i}$. It is straightforward to check that identities prevent existence of a potential function in this context, and therefore we cannot rely on this stronger concept of equilibrium.

[8]The concept of improving path is introduced by Jackson and Watts (2002), then the definition is a refinement of that concept following the seminal work by Chwe (1994) and the adaptation of the last paper to network context given by Page Jr et al. (2005).

[9]Corresponds to Definition 3, in page 531 of the paper.

[10]Corresponds to Definition 4, in page 532 of the paper.

Definition 12. *A set of networks* $G \subseteq \mathbb{G}$ *is pairwise farsightedly stable with respect to* U *and a distribution of* θ *if*

- $\forall g \in G$,
  - $\forall ij \notin g$ *such that* $g+ij \notin G$, $\exists\, g' \in F(g+ij) \cap G$ *such that* $(U_i(g', \theta_i), U_j(g', \theta_j)) = (U_i(g, \theta_i), U_j(g, \theta_j))$ *or* $U_i(g', \theta_i) < U_i(g, \theta_i)$ *or* $U_j(g', \theta_i) < U_j(g, \theta_i)$,
  - $\forall ij \in g$ *such that* $g - ij \notin G$, $\exists\, g', g'' \in F(g - ij) \cap G$ *such that* $U_i(g', \theta_i) \leq U_i(g, \theta_i)$ *and* $U_j(g'', \theta_j) \leq U_j(g, \theta_j)$,
- $\forall g' \in \mathbb{G} \setminus G$, $F(g') \cap G \neq \emptyset$
- $\nexists G' \subseteq G$ *such that* G' *satisfies previous conditions.*

Proposition 13. *Under farsightedness there are no cycles.*

Proof. See Appendix. □

Result in Proposition 13 are crucial, because we can now exploit the avoidance of cycles to explore better some applications. Moreover, among the main implication, I think is particularly relevant the role of peripheral agents. They are in fact those who will shape the equilibrium networks. The reason for that is that they would get the lowest utility in the network given a same set of links. In simple words, we could say that other agents would not like to be peripheral, and therefore they are prompt to not disappoint them. This behavior is partly driving the result of uniqueness. Moreover analyzing their behavior solely, would be quite informative if we combine the result of the two Propositions presented in this section. To make the simplest example, if we observe that the two peripheral agents of a distribution of characteristics are connected, we should expect the network to be complete. If they are both connected with the median agent the network would be surely connected, meaning that there exists an undirected path between any couple of agents.

2.4 Applications

Given the result obtained through Proposition 13, we have now a framework that allows for any comparative statics exercise on the parameters. I will do that exploring three different applications of the model. In any of these I will focus the attention on the role of some specific parameters, restricting the parameters not of interest. The first exercise refers to a threesome (restriction on N) and I will focus on conditions on types such that the possible network configurations are stable. Then I will move on a case that I called "a majority and a minority" (restriction on number of identities) showing conditions such that the two societies are ever connected, and exploring the role of parameters of benefit δ, the sensibility α, numerosity of the two societies and finally equilibrium behaviors. Finally I will move on a case where the model is interpreted in terms of peer effects in the classroom where I discuss the latest findings of the empirical literature about the "optimal policy" and the compositional treatments in classrooms. Related to that also the relation of the best reply structure with the Bonacich centrality, discussed in section 2.4.3. Then a different utility function is discussed, showing that the main characteristics of the model rely on best reply, rather than a specific parametric choice of the utility function.

2.4.1 Two Cultures: a Majority and a Minority

Through this example I wish to highlight the role of connections on equilibrium behaviors. To make the things as simple as possible I assume that there are only two cultures, and call them θ_1, θ_2, with $\theta_1 < \theta_2$. Then, we further assume that there are N agents, with $N = n_1 + n_2$ and $n_1 < n_2$. Therefore there is a majority and a minority. The main focus here is on conditions under which the two cultures would ever form a link between them.

We could interpret this situation with two cultures in same physical space, such as atheist and religious, or oriental and western cultures, or black and white. Therefore the results are interpreted as the conditions for the two cultures to be connected, which would translate into a melt down of cultures. The first thing we observe is that forming connection with agents belonging to the same culture is costless, and therefore all agents in a culture i with $i \in \{1, 2\}$ will be connected with all the $n_i - 1$ agents that share the same cultural traits. This is an application of Remark 10. The scenario would be then

$$\begin{aligned} x_i^* &= \theta_i \\ U_i &= \delta(n_i - 1) \end{aligned} \tag{2.4}$$

for all $i \in \{1, 2\}$. Now we consider adding an extra link for all players toward the rival culture, and we provide conditions on "how distant" the two cultures must be in order to be connected. The first step in addressing this point is to consider the equilibrium behaviors. The more natural way to proceed would be to evaluate equilibrium behaviors under the two scenario, and then analyze the differential in utility before and after the addition of the link. However this is not straightforward, since this would be equivalent in evaluating the marginal utility of adding a further link from the scenario described above, but utilities are discrete in links, and non linear in all other terms.

To overcome these difficulties I study behaviors considering the size of the majority group going to infinity. Then the assumption of $n_1 < n_2$ would translate in n_2 approaching the limit, while n_1 remains finite. Moreover, without loss of generality we assume that $\theta_1 = 0$, such that we can focus on the value of θ_2 under which forming links among the two cultures is profitable for both. Next Proposition shows the results under these assumption, the proof explains the mechanisms behind, and the this section is concluded discussing possible implications.

Proposition 14. *For all δ, α, n_1, there exists a threshold $\tilde{\theta_2} = \xi(\delta, \alpha, n_1)$ above which the majority would form links, but the minority refuses to.*
When links are refused, $x_i^* = \theta_i$, for $i = \{1, 2\}$.
When links are accepted $x_2^* = \theta_2$, while $x_1^* = \frac{\alpha}{n_1 - \alpha(n_1 - 1)}\theta_2$.

Proof. See Appendix. □

The intuition is that when a group is large, adding a link would influence mildly the average behavior of the group with an associated culture. On the contrary, when the group is small adding one link toward a different culture would modify consistently average behavior of agents' reference group and therefore in the former case agents are willing to do so, while in the latter they do not.

Remark 15. *The threshold $\xi(\delta, \alpha, n_1)$ is increasing in δ and n_1 for all values of α. The effects of α are non monotonic.*

To provide visual representation of the behavior of the threshold $\xi(\delta, \alpha, n_1)$, I show through the next figure a plot of the threshold as a function of α, considering different values of n_1. From this result we see the interesting implication that it could be the case, in seeking cohesion, that a country population is more than willing to welcome immigrants, but those would refuse to form such connections, harming the cohesion in the first place. The counter-intuitive result, and policy implication would be to provide incentives to minorities in participating to initiatives of the hosting country. As a parallel prediction, cohesion would be easier to achieve if the size of the minority is relatively big. As an example we may think to the Turkish community in Germany. They are a large minority and they successfully integrated into the German culture, even though they were quite diverse under several dimensions.

Finally, we saw as a result in Proposition 14 that when the link takes place agents in the minority are influenced by the hosting culture. We could interpret that as the "melting pot", meaning that different cultures tend to melt down into a unique one, over time. In an overlapping generation scenario we would then get that there will be

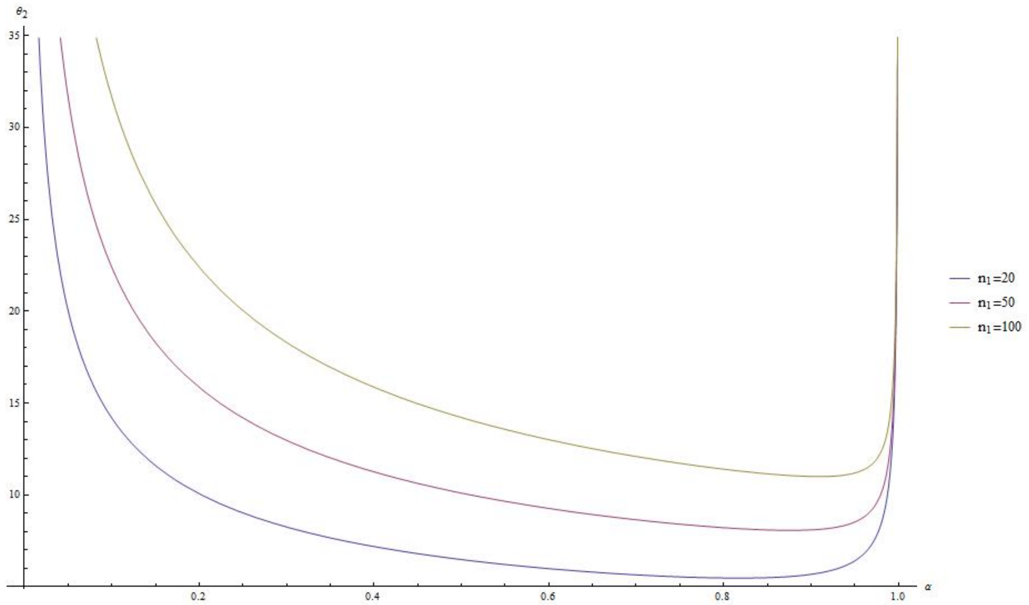


Figure 2.2: This figure shows the threshold values of θ_2 above which forming a connection is not profitable for the minority. It was assumed $\delta = 1$. Intuitively these values depend on α, represented on the horizontal axis. The three curves show from the bottom values of n_1 respectively of 20, 50 and 100. We see how the greater the minority the greater the propensity to form a link with the majority, for given values of θ_2.

convergence toward a unique culture, with characteristics of both the majority and the minority, only if the two cultures are linked in the first place. Otherwise the minority would remain segregated over time.

2.4.2 Peer Effects in the Classroom

This example attempts to provide some insights on the effects of endogenous network on peer effects. In particular, the literature on peer effects has focused recently on compositional treatments, meaning that experimentalists have changed the composition of the classrooms in terms of previously observed outcomes of the students. For instance there is the paper by Carrell et al. (2013) which in the first place states how peer effects affect heterogeneously students according to their grade; Low achievement students (LAS, from now on) are those who benefit more from the exposition to high achievement students (HAS) in a fully randomized classroom. Provided that, the authors develop a linear programming optimization whose outcome attempts to be the best composition of students' ability according to the heterogeneous effects detected in the first place. As a result the algorithm generated a partition of classes such that roughly half of them where composed by all medium achievement students (MAS), and the others by LAS and HAS. The latter are defined as bimodal classrooms. Peer effect turned out not only to be lower than expected, but even lower in the comparison with the control group, composed by randomized classrooms. Post experimental evidence showed that this was due to lack of interactions. In other words, agents formed endogenously a network composed by two segregated components.

In order to let the model being able to provide interpretation on these dimensions, I introduce some modifications to the version proposed in the main section. First of all we let the action x_i represents a student grade. This may be resulting from student's ability, with more able students having a lower cost for effort, which would justify the result that more able students are also more incline in exerting effort. However I think to that as a more complicated process, that could derive from a combination of characteristics, including work values, family background and so on. Therefore we could think to the type as $\theta_i = \sum_{k=1}^{K} \lambda_k \theta_{i,k}$, with λ_k representing a weight to any characteristic, and we let these weights be homogeneous among population.[11] For this reason this would be an extreme reduced form of the true data generating process, but we can study the network formation. Yet, assortative matching over grades is a relevant dimension to explore as documented by Barnes et al. (2014).

Moreover in the linking process it is reasonable to think that agents may have a preferential attachment over higher types, as they could strategically benefit from interacting with the smarter students. To accomplish that we could simply let the benefit be an increasing function over types. Therefore $\delta = f(\theta)$ with $f' > 0$.[12] We do not need assumptions on the curvature of this function, as it would only affect thresholds, without harming the mechanics behind the results. In this environment we have not defined yet how to measure peer effects. Next Definition formalizes it.

Definition 16. *In this context peer effects are measured as the difference of equilibrium behaviour with connections and equilibrium behaviour in isolation. Denoting peer effects as π_i we have that $\pi_i = x_i^\star(g_i') - x_i^\star(g_i^\circ)$, where g_i° is the empty network of agent i and g_i' is every other network different from g_i°.*

Before going into the analysis of this twist of the model, I would spend few words justifying this application. To further convince that the model can be accommodated to

[11] Although would be probably more interesting, letting the weights be individual specific would substantially complicate the analysis. Considering a more general framework with multi-dimensional types is left for future work.

[12] Note that the analysis here would be robust to the linear structure of the benefit assumed so far. Nevertheless, I believe this is a more realistic representation of facts, and also reduces the effect of negative peer effects, for which there is low evidence in the literature.

explain peer effects we show that the best replies structure can in fact be considered as a credible data generating process for peer effects. As a first step, recall the best reply functions. Given that in equilibrium all agents would respond to this best reply, we can rewrite 2.3 as follows

$$x_i = \alpha \frac{1}{d(i)} \sum_{j \in N(i)} (\alpha \{\bar{x_k}\}_{k \in N(j)} + (1-\alpha)\theta_j) + (1-\alpha)\theta_i$$

Now assume that all agents share the same set of peers, meaning that we are in the complete network. Therefore all average values are defined for a size of population of $N(i) = n - 1$. Collecting terms we can write it as

$$x_i = \alpha^2 \bar{x}_{N(i)} + \alpha(1-\alpha)\{\bar{\theta_j}\}_{j \in N(i)} + (1-\alpha)\theta_i$$

where $\bar{x}_{N(i)} = \frac{1}{d(i)} \sum_{j \in N(i)} \{\bar{x_k}\}_{k \in N(j)}$. Note that it is exactly the linear in means model of Manski (1993). In fact, taking the expected value of both sides we have that

$$E(x) = \alpha^2 E(x) + \alpha(1-\alpha)E(\theta) + (1-\alpha)\theta_i + \tilde{\varepsilon}_g$$

from which we get

$$E(x) = \beta_1 \theta_i + \beta_2 E(\theta) + \tilde{\varepsilon}_g$$

where $\beta_1 = \frac{1-\alpha}{1-\alpha^2}$ and $\beta_2 = \frac{\alpha}{1+\alpha}$.

It is worthy to note that $\bar{x}_{N(i)} \neq \bar{x}_{j-i}$, and $\bar{x}_{N(i)} = \bar{x}_{j-i}$ if and only if we are in the complete network. Otherwise averaging the values would wash out the heterogeneity in reference groups, which is a key source of information in this setting.

I have shown in previous sections that there exist thresholds on the "distance" on types such that agents will not form a link, because too costly. Nevertheless, let us see the distributions of types used in Carrell et al. (2013).

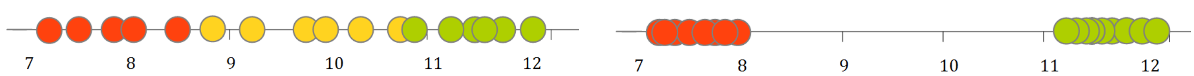


Figure 2.3: The figure shows two suggestions of possible distribution of students under the compositions assigned in Carrell et al. (2013). This is not a rigorous representation. On the axis I measure grades, from 7 to 12.

The scope of this figure is purely to suggest the possible mechanics that took place in the experiment mentioned above. The careful reader should already have understood the direction I am willing to undertake in this section of the paper. Next Remark should definitely wipe out any doubt on that.

Remark 17. *To have peer effects for LAS the network must be connected. Under segregation LAS are not positively influenced by other agents.*

Note that the assumptions on the benefit increasing with θ would yield here, as a result, that HAS gather together always, as almost all other students are willing to form links with them. Nevertheless, HAS would refuse links to LAS. If the distribution is however smooth enough, connectivity can be achieved, and peer effects would flow into the network. Since HAS are strongly connected among them, they would not be harmed by negative influence, while LAS would benefit from indirect exposure to better students.

Remark 18. *If HAS are willing to form links with LAS, the latter are willing to. The contrary is not necessarily true.*

The last remark is a weak result, that follows directly from the results above and the further assumption on non-linearity in the benefit from interactions. Therefore a formal proof is omitted. The first point in the result could be motivated through the paper by Li et al. (2014). In that paper authors run a field experiment where they ask HAS to interact with LAS. To foster these interactions experimenters paired HAS with LAS in the same desk, and monetary incentives were provided, conditional on achievement of LAS. The results were successful. A different treatment provided such incentives to only LAS, but they were not able to improve their schooling outcomes, with respect to the control group, to whom no incentives were provided. I interpret this as clear evidence of the effectiveness of the model in the context of peer effects.

The further implication of the result could be drawn integrating the evidence with the paper by Carrell et al. (2013). The authors found out that the treatment with bimodal classrooms showed less positive influence for LAS than in the randomized, control group classrooms. Post experimental evidence showed also that in fact segregation arouse in bimodal classrooms, with HAS interacting among them, and the LAS left therefore behind in homework tasks. As occasionally happens, this was an example of how the *natural variation* dominated the *optimal policy*, if the latter was designed without much theoretical understanding of the mechanics behind both the observed results, and those one is willing to obtain.

I believe that similar considerations could be applied to other relevant contexts. For instance I could mention the experiment by Kling et al. (2007). In this paper, which is part of a longer series of works, they are shown results on so called neighborhood effects. The authors have performed an experiment where they could move some individual from distressed neighborhoods to richer areas of the city, randomizing the participation into the experiment. As a result, those who moved showed no increase in any economic and psychological outcome. This was interpreted as non-relevance of neighborhood effects. In the light of the results shown here, this would rather be a consequence of segregation, arouse becuase of the endogenous network formation.

2.4.3 Nash Equilibrium and Bonacich Centrality

Recent papers show that peer effects are conceptually connected with Bonacich centrality measures. Since in this environment we are actually studying peer effects, we should expect that here also results can be expressed in terms of Bonacich centrality. Although, the interpretation of it is less informative than it is, for instance, in Ballester et al. (2006), because the position into the network, and consequently the Bonacich centrality, is endogenous. On the other hand we can show that, since agents can choose location into the network, they will have in equilibrium different preferences on how much to be central, in one sense. Contrarily to other findings, here there is not a monotonic relation between centrality and strength of peer effects, but it is non linear, and it depends on the distribution of types.

First of all, we have to find an explicit relation of best replies and Bonacich centrality. To do that, we can represent in compact form the system of best replies of the agents, that is as following.

$$\mathbf{x} = \alpha \mathbf{W} \mathbf{x} + (1-\alpha)\mathbf{t} \tag{2.5}$$

where $\mathbf{x}$ is the vector of best replies, $\mathbf{W}$ is the weighted adjacency matrix, where the weights are equal to $1/d_i(g)$, and they are different for every agent, meaning that every row is made of the same coefficients that sum up to one, but every row may differ, since agents will choose differently number of links. Finally, $\mathbf{t}$ is the vector of types. Given that, it is easy to derive the relation with Bonacich centrality. Results are summarized in the next remarks, and derivation of such result is reported below.

Remark 19. *Best replies can be expressed as a linear combination of the Bonacich centrality, and they are equal to*

$$\boldsymbol{x} = (\boldsymbol{I} - \alpha\, \boldsymbol{W})^{-1}(\boldsymbol{I} + (1-\alpha)\boldsymbol{t}) \tag{2.6}$$

where $(\boldsymbol{I} - \alpha\, \boldsymbol{W})^{-1}$ *is the Bonacich centrality measure.*

Proof. See Appendix. □

Given this simple result we end up with an equation composed only by observable parameters. Therefore we could bring this equation to the data, and study, for instance the relevant characteristics for a given population (e.g. students) that agents take into account when deciding to form a link. This sort of analysis could contribute to the paper by Marmaros and Sacerdote (2006), in which the authors could explain the relevance of characteristics behind friendships formation processes, through a unique data set and experimental context. Here instead the econometrician would only need information on the network to retrieve Bonacich centrality, and a bunch of explanatory variables to test which are the relevant ones.

2.5 Discussion

We have seen a framework of endogenous network formation taking into account optimal behaviors of agents simultaneously. Payoff depend in particular on identities of agents and social influence that they would get under a feasible network configuration. Agents would form the network exhibiting homophilous behavior, but at the same time there is some degree of diversity they are able to accept. In other words, identities can be interpreted as optimal behavior prior to the interaction with other players, while the outcome would be the *ex-post* behavior of agents. In these terms, the identity would be a reduced form of more complicated process able to determine.

When we observe segregation, one of the key implications is on which of the two parties involved in the choice of the link may refuse it. This component is particularly relevant in policy experiments that aim to target specific branches of population. In the paper I have discussed mainly the case of peer effects in the classroom. This is relevant, since it is widely recognized that compositional treatment may be able to improve the condition of targeted group of students. Policies of this kind, moreover, allow for interventions that keep the amount of resources constant, and therefore do no call for additional costs. Thus I showed how we should take care of the right design of incentives to make such policies effective. Another example consistent with the model is provided by the Moving to Opportunity experiment.[13] Recent evidence provided by Chetty et al. (2015), shows that segregation patterns were relevant when exposing agents to neighborhoods with different characteristics. Thus a proper incentive scheme may have increased the likelihood of matches between subjects endowed with different socio-economic backgrounds. In such a

[13]See Kling et al. (2007).

way the conclusions of that experiment may be reverted.

Among the assumptions that I have used to obtain results, there is perfect information. I believe that it could be extremely interesting to let identities be private information. A dynamic setting would then give room for players to strategically choose the current network to falsify their identity, if it would ever be profitable for them. In this version of the model this cannot happen since the game is static, and therefore agents have all interests in revealing truthfully their identities. However extending the setting to a dynamic one, with imperfect information would be both challenging and relevant in unveiling several real world mechanisms.

Chapter 3

Search over the network

Agents in each period choose how much time to spend on-line and off-line. Off-line life delivers utility conditional on having found nodes active. On-line activity instead is characterized by a benchmark level of utility proportional to existing degree on the virtual platform, plus a level of utility increasing with number of agents active on the platform. New links are assumed to be formed only in real life, and only a proportion of them is transferred to the virtual network too. Choices are represented through optimal stopping rules. As results agents with low probability of meetings are more prone to use virtual social networks, while agents living in dynamic and rich environments will tend to spend more time in the real life.

3.1 Introduction

Social Networks have spread around the globe surprisingly. Users also uses these tools quite often. This simple observation raises a point: there is a real network and a virtual network. The point of this paper is to explore why and when people log in into their virtual life, which implies that they log out from the real one. Agents spend a portion t of their time in the real life, and the rest of the time $1 - t$ is spent in the virtual life. Links that took place in the real life are automatically transferred as links in the virtual life, while the opposite is not necessarily true. However in real life the network is local, meaning that agents have geographical constraints. This can be represented either with a cost of meeting a link which depends on distance, or through a probability of meeting an active link, again decreasing with distance. This clearly does not apply to the virtual network.

An interesting observation that naturally arises is that if agents have a probability of meeting a link in one of the two networks, this will depend also on the choice of agents of being logged in or logged off. Therefore we can find the optimal choice for t for the agents, assuming some preferences for having some active contacts.[1] It will depend on the degree of agents and on the choices of other agents. Clearly there are two interesting equilibria, one where all agents spend lot of time in real life, and one other where agents are on-line most of their time.

It is interesting to investigate conditions on the structure of probabilities of meetings for which such states can coexist. It is indeed well documented that people uses social network differently. For some it is only a mean to enhance real connections, for other it is truly a substitute, and develop their existing social relationship mainly on-line sacrificing their real relationships.

Possibly a way to achieve that is to enrich the model with a dynamic evolution of the network. If we imagine a friend of friend type of network as in Jackson and Rogers (2007), we should have that there are basically two equilibria of the system: one where agents use the virtual network as an ancillary device, and spend most of their time on the real network. The reinforcement comes from the fact that agents in this type of equilibrium are more likely to meet new agents of the same kind. On the other hand agents with low degree are stuck in an equilibrium where they spend lot of time logged in, and therefore

[1] We can think of it as living experiences, which can happen only if two or more agents have an active link in one of the two networks.

they miss the chance to meet people in the real life.

The model can be summarized as follows. Agents in each period choose how much time to spend on-line and off-line. Off-line life delivers utility conditional on having found nodes active. On-line activity instead is characterized by a benchmark level of utility proportional to existing degree on the virtual platform, plus a level of utility increasing with number of agents active on the platform. New links are assumed to be formed only in real life, and only a proportion of them is transferred to the virtual network too. This is to reflect that not all individuals use the platform, or at least not the same, although we assume here that there is only one virtual platform.

The analysis is on two dimensions, within period and between periods. The process is largely history dependent, so we will bring comparative statics results mainly on two parameters of interest, the probability of meetings, which is individual specific, and the proportion of adoption of the virtual platform which is global. Agents will thus choose a platform depending on the activity of other players. In the beginning of the period agents will spend mostly their time off-line, so to form social contacts. Once their level of degree in virtual life is high enough agents will begin considering to spend an increasing fraction of time on-line. However having real life activities is always desirable since it would represents an investment on future utilities.

3.2 Literature review

In general the paper contributes to the young literature on multiplexing in networks. By that we mean that networks interact with each other. Thus, choices on one network will determine spillover effects on other layer of the same network. Little is known in economics about this issue. To the best of my knowledge there are no papers discussing this issue, yet.[2] We know a substantial amount of results for single-layered networks, but how a multi-layer analysis may change existing results is yet to be determined. In principles it complicates the analysis, but it will surely deliver a better understanding of the reality.

Another stream of literature, purely economic, is about search models. Such literature was initialized during '70s by the paper McCall (1970) which brought the knowledge on optimal stopping rules on economic decisions, mainly in job search. For the methodologies the paper here reminds the stream of literature associated to that paper, although the

[2] A remarkable example is still a work in progress by Banerjee et al. (2016).

process here differs for several reasons. First of all the search here is a cumulative process of nodes, while those paper analyze independent job offers that arrive exogenously. Here instead agents will search for nodes, and their choices are interdependent.

Another literature related regards the topic of virtual social networks. Such thing is deeply studied in fields other than economics, such as psychology, computer science, physics and political science. Some of the insights on overall research can be found in Wilson et al. (2012), which provide a broad overview of all articles that have been written on Facebook and similar issues. Here we contribute to these papers, too, because we aim to shed light on some of the mechanism of adoption and some of the observed patterns of consumption of on-line social networks.

The rest of the paper is organized as follows. First we discuss the model and the assumptions behind it. Thus we describe the process of search over the network and how to solve the model.

3.3 The Model

We assume that there is an infinite number of time steps S of length T. Time within blocks is discrete and indexed by $t \in \{1, 2, ..., T\}$ with a finite horizon $T < \infty$. Agents get utility from a function f of their number of friends met. This function is increasing in the degree d and the amount of time available $(T - t)$. For the time being we do not need any assumption on the second derivatives of this function. Now we analyze behavior within time step S, so we drop the subscript to not abuse with notation. Thus utility of an agent i at time t is given by the following

$$\begin{aligned} U_{it} &= f(d_i, t) + \mathbb{1} g(\alpha d_i, t) \\ &= \sum_t (f(d_i) + g(\alpha d_i)) \end{aligned} \tag{3.1}$$

where α denotes the fraction of user in the population that have adopted the social network, and $\mathbb{1}$ is an indicator function assuming value 1 if adopted, 0 otherwise. We assume that the representative agent we solve for has adopted the social network. Alternatively agents can spend their time on a virtual social network, which yields a fixed level of utility $V(t)$ which depends only on time. The available time to agents can also be invested in a search activity, where they are intended to organize an event. They can do that both on-line and off-line. This choice is made at any time step depending on the expected probability of meeting in real world, and meeting on the social network. This probability is an outcome of the coordination game. In principle agents would like to coordinate on a search platform. The following set of assumptions together with the preferences explained before, fully describe the model.

Definition 20. *An agent i continues searching whenever $f(d_i, t) \leq EU_{it+1}$ and $EU_{it+1} \geq g(\alpha d_i)$.*

So any agent at the beginning of the period they have to choose whether to search or not. Something that follows immediately from the definitions of the model is that the probability of meeting should be high enough, or the reward for meeting, measured with the function f should be high enough, otherwise search is not at all profitable, and agents may prefer to enjoy their reservation utility for the whole period. In other words, investing in search is not profitable in expectation, so agents may not engage in search at all.

3.4 Search over the Network

The meetings are random and we denote them with $p(m)$. Preferences are over degree and time spent with agents. The idea is that there repeated windows of time where agents perform the search and after it they spend the remaining time with the nodes they found. At the end of the period the process start again.

Note that we are assuming that there is no benefit spending time alone, i.e. the search period does not pay off the agents any utility. Nevertheless this is without loss of generality since we could assume a benefit deriving from the search without altering the second order effects of the optimization.[3]

At any point in time t an agent faces the choice of certain versus uncertain. She knows the utility she would get stopping the search, and formulate rational expectations over the utility of continuing the search. Here we further assume that the meeting occurs only if both agents are searching. This is to introduce a complementarity between search. To see this let me show the optimal stopping rule.

The process is defined such that the search continues until the martingale condition

[3]This would be the equivalent of the unemployment benefit b used in most of the models of job search.

is met. Given the preferences in equation 3.1 we have that

$$\begin{aligned} U_{it} &\leq EU_{it+1} \\ (T-t)f(d_{it}) &\leq [f(d_{it}) + p(m)p_{-i}(s)](T-t-1) \\ f(d_{it}) &\leq p(m)p_{-i}(s)(T-t-1) \end{aligned} \tag{3.2}$$

and therefore solving for t we obtain

$$t_i^* = T - \frac{f(d_{it})}{p(m)p_{-i}(s)} - 1 \tag{3.3}$$

As interesting remarks we see that the search duration depends positively both on the probability of meetings and on the probability of other agents' search time. These two terms are not explicit yet, but we could see how the process changes assuming different random meetings models. That would yield a closed form solution to equation 3.3.

3.4.1 Expected degree

In the first period we can calculate the expected degree of agents. This is given in fact by the probability of finding a node through the search, and this is given by p. Then every node has a certain probability to be found by other nodes, through their search. However, he can be searched by all agents if the link formation with the first agent is not successful. This happens with probability $1-p$. If instead the direct connection is successful the agent cannot be found by the agent with which she is already linked with. Therefore all together the expected degree of an agent i in the first period is given by

$$\begin{aligned} E(d_{i1}) &= p + p^2\frac{N-2}{N-1} + p(1-p) \\ &= p(2 + p\frac{N-2}{N-1} - p) \\ &= p(2 - \frac{p}{N-1}) \end{aligned} \tag{3.4}$$

What the previous formula says is that agent i at period 1, can pick an agent through her search, and the probability to do that in expectation is the average probability of meeting $E_i(p_{ij}) = \frac{1}{N-1}\sum_{j\in N\setminus i} p_{ij}$. However i can also be found through other agent's search activity. Thus if i found a node, there are other $N-2$ nodes available. They will find agent i with a given probability, which is weighted down by $N-1$ since they can still find any other agent rather than i. This happens with probability $E_i(p_{ij})$. With the remaining probability, agent i has not found anybody, so the full set of agents can find him.

This formula however is non general, so we can see a more general one below, denoting with $E_i(p_{ij}) = \frac{1}{N-1} \sum_{j \in N \setminus i} p_{ij}$

$$E(d_{i1}) = E_i(p_{ij}) + E_i(p_{ij}) \sum_{k \in N \setminus i,j} \frac{p_{ki}}{N-1} + (1 - E_i(p_{ij})) \sum_{j \in N \setminus i} \frac{p_{ji}}{N-1} = \\ = E_i(p_{ij}) + E_i(p_{ij}) \sum_{k \in N \setminus i,j} \frac{p_{ki}}{N-1} + (1 - E_i(p_{ij}))E_i(p_{ji}) \tag{3.5}$$

For large N, we can approximate $\sum_{k \in N \setminus i,j} \frac{p_{ki}}{N-1} \approx E_i(p_{ji})$, such that 3.5 can be rewritten as

$$E(d_{i1}) = E_i(p_{ij}) + E_i(p_{ji}) \tag{3.6}$$

In the non general case where $p_{ij} = p$ for all i, j, we would simply have $E(d_{i1}) = 2p$. Iterating the formula we can calculate the expected degree in any period $t + 1$, which depends on the expected degree in period t, since this is a cumulative process.

$$E(d_{it+1}) = E(d_{it}) + p(2 - \frac{p}{N - E(d_{it}) - 1}) \tag{3.7}$$

However, we have still to take into account the agents that may have left the search process, because that would enter in the denominator of the second term in the right hand side of equation 3.7.

The general formula for the optimal stopping time t^*, given a degree d_{it^*} is

$$E(t_i^* | d_{it^*}) = T - \frac{d_{it^*}}{p(2 - \frac{p}{N - d_{it^*} - 1})} - 1 \tag{3.8}$$

From this we can determine the degree such that the optimal stopping rule holds, as a function of t. As a result it is uniquely determined for each t.[4] Call $\bar{d_{it^*}}$ the degree that allows equation 3.8 to hold true. We have now a sequence of $\bar{d_{it^*}}$, since for every t it is identified one value for the degree.

Now the process is binomial (is it?). Therefore we can calculate thanks to the sequence of $\bar{d}_{it^*}$, the probability that an agent is still searching, that is the probability that an agent would have a degree higher than $\bar{d}_{it^*}$, in any given t. We represent it through the cumulative distribution function of degrees.

[4]Being precise, solving for d_{it^*} yields two solutions. Nevertheless we are interested only in the first solution, since the other is not consistent with the model.

Thanks to the binomial formulas we get

$$F(stop|t) = 1 - \sum_{k=1}^{\bar{d}_{it^*}} \binom{t(N-1)}{k} p^k (1-p)^{t(N-1)-k} \tag{3.9}$$

For the whole sequence of time, we will have a vector of probabilities that is roughly 0 in the beginning, since all agents will be searching. Over time though, the probability that an agent stops searching increases, as well as the $\bar{d}_{it^*}$ decreases over time. This yields a cumulative distribution function of the agents that have stopped searching.

3.5 A routine for the meeting process

First we have to characterize the network depending on the two relevant dimensions explored in this paper, which is randomness and homophily.

The network formation stage is fairly general, and the main drivers for the different configurations that arise are described by two parameters, that represent respectively the nature of meeting, i.e. how *random* they are, and the tolerance of agents with respect to diversity, i.e. a measure of *homophily*. Particular combinations of these parameters let the process reconcile the standard random graph theory (Erdös and Rényi (1960)), the friend-of-friends type of networks as in Jackson and Rogers (2007), and other existing models of percolation (see DSouza and Nagler (2015) for a review of such models).

Each agent is endowed with a vector of m characteristics, and we call it $\theta \in \mathbb{R}^m$. When considering a link we let each agent compare her vector of characteristics with the agent who may potentially be her friend (a *link*). Such a rational rule would generate a score that agent i uses to evaluate whether the link is worth to be formed or not. Calling such vector θ_i and θ_j respectively, we could generally say that the score would be given by $f(\theta_i - \theta_j)$, where f denotes a generic function, on which for the moment with put no restrictions, for the sake of generality. Agents will then pick a link from a set of potential links if $f(\theta_i - \theta_j) \leq \alpha$, with α denoting the threshold of tolerance for differences in characteristics.[5]

Lot of the process would depend on the meeting structure. To represent that, I analyze

[5] Although several functions are considered, I developed examples using a specific function f, given simply by the absolute value. The rule of selection therefore becomes $|\theta_i - \theta_j| \leq \alpha$. However we could accommodate the model to exhibit different preference structures, depending on the application of interest. The rule assumed here though, remains the more interesting in this case, because it would matter how homophilous networks ware.

a process that is flexible enough to embed both randomness and friend-of-friends type of meetings. To achieve that we could describe the process according to the following routine.

1. pick an agent
2. then pick at random a set of k links
   (a) c/k are selected at random with a given probability p
   (b) $(k-c)/k$ are selected from the set of indirect links. Thus, if $\sum_j d_j > k - c$ then all the $k - c$ are picked. If $\sum_j d_j \leq k - c$ then the difference of nodes are picked at random.[6]
3. the agent chooses best link to form, i.e. $\arg\min_{\theta_j} f(\theta_i - \theta_j) \leq \alpha$
4. iterate

Such a process is flexible enough to embed the classical Erdos-Renyi random network model,[7] if we set $c = k$. On the contrary, if $c = 0$ the meetings will happen only through friends-of-friends a-la Jackson and Rogers (2007), and for intermediate values of c we would have a hybrid scenario.

The model is then analyzed assuming heterogeneity in the two key parameters, namely c and α. Those will map into a network configuration, which in turns shapes the choice in equilibrium of the intensity of the activity. This implies to solve a coordination problem, where all agents simultaneously choose the intensity x depending on other agents' choices. Note that while x represents a key dimension, but at the same time other statistics of the network topology will provide additional insights on the nature of meetings the agents have, and on the network characteristics that would bring to the several possible outcomes.

[6] Note that such a rule would allow to initiate a network composed in the first period by N empty nodes. Thus, in the beginning the first meetings happen at random, and over time the bias would gain importance until it would completely wipe out the randomness.

[7] See Erdös and Rényi (1960).

Conclusions

Through this thesis I showed how the process of formation of connection matters for social influence processes. I showed that in two ways: a context where agents have full decisional power on which connections to form, and a context where meetings follow a stochastic process, still with preferences that matter. I stressed the importance of these concepts, because this is a missing feature of current works on the field.

One interesting feature is to bring all the results proposed here to some empirical investigation. All theoretical implication have the potential to be tested on data. Clearly how successful would be the empirical analysis would depend on quality of data, but this is clearly an important direction of my research agenda.

Moreover I believe that important extensions may follow from the current state of these three papers. We could both relax some of the assumptions to make the models better fit given applications, or extend adding more structure. An example of that is to have some strategic agents that can interact with the network and that have specific preferences on the network structure. Therefore it would be interesting to characterize their behavior and study the change in outcomes, given this interference.

Bibliography

Akerlof, G. A. and R. E. Kranton (2000). Economics and identity. *The Quarterly Journal of Economics 115*(3), 715–753.

Andris, C., D. Lee, M. J. Hamilton, M. Martino, C. E. Gunning, and J. A. Selden (2015). The rise of partisanship and super-cooperators in the us house of representatives.

Ballester, C., A. Calvó-Armengol, and Y. Zenou (2006). Who's who in networks. wanted: the key player. *Econometrica 74*(5), 1403–1417.

Banerjee, A., A. G. Chandrasekhar, E. Duflo, and M. O. Jackson (2016). Changes in social network structure in response to exposure to formal credit markets. *mimeo*.

Barnes, J., K. M. Beaver, J. T. Young, and M. TenEyck (2014). A behavior genetic analysis of the tendency for youth to associate according to gpa. *Social Networks 38*, 41–49.

Bisin, A., U. Horst, and O. Özgür (2006). Rational expectations equilibria of economies with local interactions. *Journal of Economic Theory 127*(1), 74–116.

Bolletta, U. (2015). Endogenous network with identities and social influence. *Available at SSRN 2581373*.

Boucher, V. (2015). Structural homophily. *International Economic Review 56*(1), 235–264.

Bramoullé, Y. and R. Kranton (2007). Public goods in networks. *Journal of Economic Theory 135*(1), 478–494.

Bramoullé, Y., R. Kranton, and M. D'Amours (2014). Strategic interaction and networks. *The American Economic Review 104*(3), 898–930.

Brualdi, R. A. and H. J. Ryser (1991). *Combinatorial matrix theory*, Volume 39. Cambridge University Press.

Calvó-Armengol, A. and R. İlkılıç (2009). Pairwise-stability and nash equilibria in network formation. *International Journal of Game Theory 38*(1), 51–79.

Calvó-Armengol, A., E. Patacchini, and Y. Zenou (2009). Peer effects and social networks in education. *The Review of Economic Studies 76*(4), 1239–1267.

Carrell, S. E., B. I. Sacerdote, and J. E. West (2013). From natural variation to optimal policy? the importance of endogenous peer group formation. *Econometrica 81*(3), 855–882.

Chetty, R., N. Hendren, and L. F. Katz (2015). The effects of exposure to better neighborhoods on children: New evidence from the moving to opportunity experiment. Technical report, National Bureau of Economic Research.

Chwe, M. S.-Y. (1994). Farsighted coalitional stability. *Journal of Economic Theory 63*(2), 299–325.

Currarini, S., M. O. Jackson, and P. Pin (2009). An economic model of friendship: Homophily, minorities, and segregation. *Econometrica 77*(4), 1003–1045.

DeGroot, M. H. (1974). Reaching a consensus. *Journal of the American Statistical Association 69*(345), 118–121.

DeMarzo, P. M., D. Vayanos, and J. Zwiebel (2003). Persuasion bias, social influence, and unidimensional opinions. *The Quarterly Journal of Economics 118*(3), 909–968.

DSouza, R. M. and J. Nagler (2015). Anomalous critical and supercritical phenomena in explosive percolation. *Nature Physics 11*(7), 531–538.

Erdös, P. and A. Rényi (1960). On the evolution of random graphs. *Publ. Math. Inst. Hung. Acad. Sci 5*, 17–61.

Galeotti, A. and S. Goyal (2010). The law of the few. *The American Economic Review*, 1468–1492.

Gentzkow, M. and J. M. Shapiro (2010). What drives media slant? evidence from us daily newspapers. *Econometrica 78*(1), 35–71.

Golub, B. and M. O. Jackson (2010). Naïve learning in social networks and the wisdom of crowds. *American Economic Journal: Microeconomics 2*(1), 112–49.

Golub, B. and M. O. Jackson (2012). How homophily affects the speed of learning and best-response dynamics. *The Quarterly Journal of Economics 127*(3), 1287–1338.

Hegselmann, R. and U. Krause (2002). Opinion dynamics and bounded confidence models, analysis, and simulation. *Journal of Artificial Societies and Social Simulation 5*(3).

Herings, P., A. Mauleon, and V. Vannetelbosch (2009). Farsightedly stable networks. *Games and Economic Behavior 67*(2), 526–541.

Jackson, M. O. (2008). *Social and economic networks*.

Jackson, M. O. and B. W. Rogers (2007). Meeting strangers and friends of friends: How random are social networks? *The American economic review*, 890–915.

Jackson, M. O. and A. Watts (2002). The evolution of social and economic networks. *Journal of Economic Theory 106*(2), 265–295.

Jackson, M. O. and A. Wolinsky (1996). A strategic model of social and economic networks. *Journal of Economic Theory 71*(1), 44–74.

Kinateder, M. and L. P. Merlino. Public goods in endogenous networks. *Unpublished Manuscript*.

Kling, J. R., J. B. Liebman, and L. F. Katz (2007). Experimental analysis of neighborhood effects. *Econometrica 75*(1), 83–119.

Krause, U. (2000). A discrete nonlinear and non-autonomous model of consensus formation. *Communications in difference equations*, 227–236.

Li, T., L. Han, L. Zhang, and S. Rozelle (2014). Encouraging classroom peer interactions: Evidence from chinese migrant schools. *Journal of Public Economics*.

Manski, C. F. (1993). Identification of endogenous social effects: The reflection problem. *The review of economic studies 60*(3), 531–542.

Marmaros, D. and B. Sacerdote (2006). How do friendships form? *The Quarterly Journal of Economics 121*(1), 79–119.

McCall, J. J. (1970). Economics of information and job search. *The Quarterly Journal of Economics*, 113–126.

Melguizo, I. (2015). Endogenous homophily and the persistence of disagreement. *mimeo*.

Monderer, D. and L. S. Shapley (1996). Potential games. *Games and economic behavior 14*(1), 124–143.

Page Jr, F. H., M. H. Wooders, and S. Kamat (2005). Networks and farsighted stability. *Journal of Economic Theory 120*(2), 257–269.

Watts, D. J. and S. H. Strogatz (1998). Collective dynamics of small-worldnetworks. *Nature 393*(6684), 440–442.

Wilson, R. E., S. D. Gosling, and L. T. Graham (2012). A review of facebook research in the social sciences. *Perspectives on psychological science 7*(3), 203–220.

Yildiz, E., A. Ozdaglar, D. Acemoglu, A. Saberi, and A. Scaglione (2013). Binary opinion dynamics with stubborn agents. *ACM Transactions on Economics and Computation 1*(4), 19.

Appendices

Appendix A

Polarization through the endogenous network

A.1 Payoff in equilibrium

From (1.3) and (1.4), the payoff in equilibrium is

$$\begin{aligned}
\pi_i &= k_i \left(V - f^2(1-f)(\mu_i - t_i)^2\right) - f \sum_{j \in d_i} (x_i^* - x_j)^2 \\
&= k_i \left(V - f(1-f)(\mu_i - t_i)^2\right) - f \sum_{j \in d_i} \left[(x_i^* - x_j)^2 + (1-f)^2(\mu_i - t_i)^2\right] \\
&= k_i \left(V - f(1-f)(\mu_i - t_i)^2\right) - f \sum_{j \in d_i} (x_j - \mu_i)^2 \\
&= k_i \left(V - f(1-f)(\mu_i - t_i)^2 - f\sigma_i^2\right) \quad , \qquad \text{(A.1)}
\end{aligned}$$

where we have called $\sigma_i^2 \equiv \frac{\sum_{j \in d_i}(x_j - \mu_i)^2}{k}$ the variance of the actions of i's neighbors.

A.2 Proof of Lemma 1

There is clearly a unique solution to the unconstrained equation (1.5), because $(I - F)$ and $(I - FD)$ are always full rank matrices.

Now suppose that the maximum element of $\vec{x}$, call it x_m is such that $x_m > 1$. Then, by (1.4), it is a convex combination of $\frac{\sum_{j \in d_m} x_j}{\ell_m}$ and t_m. But since $t_m < 1$, it must be that $\frac{\sum_{j \in d_m} x_j}{\ell_m} > x_m$, which contradicts the initial assumption.

In the same way it is impossible that the minimum element of $\vec{x}$ is less than 0.

A.3 Payoff differential

Recall the payoff in equilibrium

$$\pi_i = k_i(V - f(1-f)(\mu_i - t_i)^2 - f\sigma_i^2) \tag{A.2}$$

From which we can define the additional payoff deriving from the addition of a link with an agent j

$$\pi_{i\to j} = (k_i+1)(V - f(1-f)\left(\frac{\mu_i k_i + x_j}{k_i+1} - t_i\right)^2 - f\sigma_{i\to j}^2) \tag{A.3}$$

Given equation A.3, we can now define the payoff differential, simply given by the difference $\Delta\pi_{i\to j} = (\pi_{i\to j} - \pi_i)$

$$\Delta\pi_{i\to j} = V - f(1-f)\left((k_i+1)\theta' - k_i\theta\right) - f\left((k_i+1)\sigma_{i\to j}^2 - k_i\sigma_i^2\right) \tag{A.4}$$

where $\theta' = \left(\frac{\mu_i k_i + x_j}{k_i+1} - t_i\right)^2$ and $\theta = (\mu_i - t_i)^2$. Let us now focus on this difference. Thus

$$\begin{aligned}
&(k_i+1)\theta' - k_i\theta = \\
&= (k_i+1)\left(\frac{\mu_i k_i + x_j}{k_i+1} - t_i\right)^2 - k_i(\mu_i - t_i)^2 \\
&= (k_i+1)\left((\mu_i - t_i) - \frac{1}{k_i+1}(\mu_i - x_j)\right)^2 - k_i(\mu_i - t_i)^2 \\
&= (\mu_i - t_i)^2 + \frac{1}{k_i+1}(\mu_i - x_j)^2 - 2(\mu_i - t_i)(\mu_i - x_j) \\
&= (t_i - x_j)^2 - \frac{k_i}{k_i+1}(\mu_i - x_j)^2
\end{aligned} \tag{A.5}$$

Now let us focus on the term $\left((k_i+1)\sigma^2_{i\to j} - k_i\sigma_i^2\right)$

$$
\begin{aligned}
&\left((k_i+1)\sigma^2_{i\to j} - k_i\sigma_i^2\right) = \\
&= \frac{k_i+1}{k_i+1}\left(\sum_h (x_h-\mu_i')^2 + (x_j-\mu_i')^2\right) - \sum_h (x_h-\mu_i)^2 \\
&= \left(\sum_h \left(x_h - \frac{k_i\mu_i + x_j}{k_i+1}\right)^2 + \left(x_j - \frac{k_i\mu_i+x_j}{k_i+1}\right)^2\right) - \sum_h (x_h-\mu_i)^2 \\
&= \sum_h \left(x_h - \frac{k_i\mu_i}{k_i+1} - \frac{x_j}{k_i+1}\right)^2 - \left(\frac{k_i}{k_i+1}(x_j-\mu_i)\right)^2 - \sum_h (x_h-\mu_i)^2 \\
&= \sum_h \left(x_h - \mu_i + \frac{\mu_i}{k_i+1} - \frac{x_j}{k_i+1}\right)^2 - \frac{k_i^2}{(k_i+1)^2}(x_j-\mu_i)^2 - \sum_h (x_h-\mu_i)^2 \\
&= \sum_h (x_h-\mu_i)^2 + \sum_h \left(\frac{1}{k_i+1}(\mu_i-x_j)^2\right) - 2\sum_h\left(\frac{1}{k_i+1}(x_h-\mu_i)(\mu_i-x_j)\right) + \\
&- \frac{k_i^2}{(k_i+1)^2}(x_j-\mu_i))^2 - \sum_h (x_h-\mu_i)^2 \\
&= \frac{k_i}{k_i+1}(1-\frac{k_i}{k_i+1})(\mu_i-x_j)^2
\end{aligned}
\tag{A.6}
$$

Therefore combining the two previous results we get

$$
\begin{aligned}
\Delta\pi_{i\to j} &= V - f(1-f)\left((t_i-x_j)^2 - \frac{k_i}{k_i+1}(\mu_i-x_j)^2\right) \\
&- f\left(\frac{k_i}{k_i+1}(1-\frac{k_i}{k_i+1})(\mu_i-x_j)^2\right) \\
&= V - f(1-f)\,(t_i-x_j)^2 - f\frac{k_i}{k_i+1}\left(f - \frac{k_i}{k_i+1}\right)(\mu_i-x_j)^2
\end{aligned}
\tag{A.7}
$$

A.4 Proof of Proposition 1

Proof. We proceed by steps, first addressing existence constructing the equilibrium. In particular we show that any equilibrium configuration must be ordered, as defined in 4, which then partially characterizes the equilibrium, too. Finally we show that it is also unique.

It is easy to check that if the equilibrium is ordered we have that

$$
\begin{aligned}
x_i > x_j &\leftrightarrow t_i > t_j \\
x_i < x_j &\leftrightarrow t_i < t_j \\
x_i = x_j &\leftrightarrow t_i = t_j
\end{aligned}
\tag{A.8}
$$

Given that, we show that it cannot exist another equilibrium different from a ordered one, such that any of the above conditions hold true. Consider now the minimum x_i in the population, and call it x_0. Without loss of generality we let $x_0 \leq x_1 \leq ... \leq x_N$, and to prove the result basically we want to match every action with a type such that $t_i \mapsto x_i$ for all $i \in N$. Thus we claim that i is such that t_i is the minimum initial opinion in the population, calling it t_0. We can see this through the payoff in equilibrium, which we recall

$$k_i \left(V - f(1-f)(\mu_i - t_i)^2 - f\sigma_i^2\right)$$

Recalling the equation we see that the highest utility can be achieved minimizing the value of μ_i, given that we are considering t_0, and the variance term ensures sharp bounds on it. To see it rigorously we derive the payoff differential deriving from the addition of a link, and which is given by

$$\Delta\pi_{i\to j} = V - f(1-f)\left(t_i - x_j\right)^2 - f\frac{k_i}{k_i+1}\left(f - \frac{k_i}{k_i+1}\right)(\mu_i - x_j)^2 \qquad \text{(A.9)}$$

Therefore links are added as long as the equation above is positive. Moreover any non-ordered scenario would let profitable deviation for t_0, and thus that would not be an equilibrium. Note that this holds for any $f \in (0,1)$. Indeed if $f < \frac{k_i}{k_i+1}$ the second term on the right hand side would be positive. Nevertheless if $f < \frac{k_i}{k_i+1}$ then $f(1-f) > f\frac{k_i}{k_i+1}\left(f - \frac{k_i}{k_i+1}\right)$ is always true, and thus this does not alter the scheme we have built. If $f = 0$ all links would be added, and if instead $f = 1$ then $f > \frac{k_i}{k_i+1}$. This shows that in equilibrium the lowest t_i will exhibit the lowest x_i. Similarly this holds for x_N and t_N.

Now let us focus on the second lowest value of x, x_1. Following a similar argument, we claim that it should belong to the agent with the second lowest opinion in the population, t_1. This is slightest more tedious to prove, since agent i can form links both at her right and left. We can in fact reproduce the procedure we used to prove that $t_0 \mapsto x_0$, with the extra step set by the presence of agent t_0. Similarly with respect to agent t_0, agent t_1 would form links with agents $x_2 \leq x_3 \leq ... \leq x_{b_1}$, where b_1 is the optimal threshold as in the definition of ordered equilibrium. Moreover it is likely that agent t_1 would form a link with t_0. To see this we should see again payoff in equilibrium. The addition of such link would generate an increase in σ_1^2, but a decrease of the term $(\mu_1 - t_1)^2$. Clearly it can happen that such a link is not optimal, but in either case it would not contradict the ordered equilibrium scenario. Symmetrically this holds for t_{N-1}.

Therefore we showed that in equilibrium $t_i \mapsto x_i$ for $i \in \{0, 1, N-1, N\}$. Iterating the reasoning, with the due differences because of the availability of agents at both sides, the result holds for all other agents.

The set of relations described by A.8 is always true in equilibrium. We have shown that if the above set of condition holds the equilibrium must be ordered and that in equilibrium the ordered scenario is the only possible outcome. This thus addresses existence and the partial characterization.

This concludes the proof. □

A.5 Proof of Proposition 2

Proof. The proof is built on the intuition that if two agents would never connect together if isolated, they would not do a fortiori whenever they are surrounded by other agents. So if we consider two nodes i, j, we have it will never be profitable for them to form a connection whenever the following holds

$$|t_i - t_j| \geq \left(\frac{(1+f)^2}{f(1-f)}V\right)^{\frac{1}{2}} \tag{A.10}$$

To see that this is indeed a sufficient condition we may observe that the link will never be a profitable deviation under any possible scenario.

- We are in sub-game perfect Nash
- We want two nodes $i, i+1$ that form respectively groups such that
  1. $N_i(g)|\{i - a_i, ..., i-1\}$
  2. $N_j(g)|\{j+1, ..., j+b_j\}$
- Assume i is in that scenario, and consider a single-deviation from that equilibrium.
- If $N_i(g) \in \emptyset$, whenever $\left(\frac{(1+f)^2}{f(1-f)}V\right)^{\frac{1}{2}} < |t_i - t_{i+1}|$, then $\pi_i < 0$.
- For all $N_i(g) \notin \emptyset$ payoff are even lower given

$$\Delta\pi_{i\to j} = V - f(1-f)\left(t_i - x_j\right)^2 - f\frac{k_i}{k_i+1}\left(f - \frac{k_i}{k_i+1}\right)\left(\mu_i - x_j\right)^2$$

- It is easy to check that the same holds for j, too, because of symmetry.

Through this we have proved the first two points of the proposition. It remains the third. However it simply states that there will be a number of components $|C|$ is going to be at least equal to the number of couples of agents that do not satisfy the sufficient condition for the network to be connected, which means that it exhibits a single component. Easily if there is a single couple ij, then $|C| = 2$. It is a lower bound, since we cannot control the components that took place because of the rule of order chosen to select one of the equilibria. $\square$

Appendix B

Peer effects in schools: a model

B.1 Proof of Lemma 8

Proof. The problem is represented by a system of N equations in N unknown variables. By linearity of equations and restrictions on the parameter α the solution of the system is always unique. If $\alpha = 1$ there is a continuum of equilibria, but I am not interested in such a case because of triviality of the results, and the restriction on the parameter is assumed throughout the rest of the paper. To show that $x_i^* \in [\underline{\theta}, \bar{\theta}]$ we could use a simple contradictory argument. For instance just assume that there is an agent ℓ such that $x_\ell > \bar{\theta}$ and that it is the maximum value among all x_i^*. Given that behavior in equilibrium is a convex combination of $\{x_j\}_{j \in N_\ell}$ and θ_ℓ, as stated in equation 2.3, that $\theta_\ell \in [\underline{\theta}, \bar{\theta}]$, $x_\ell > \bar{\theta}$ if and only if $\{x_j\}_{j \in N_\ell} > x_\ell$, which contradicts the initial assumption. Same argument can be applied to show that $x_i^* \geq \underline{\theta}$. □

B.2 Proof of Proposition 9

Proof. To verify this Result we can draw agents as ordered on a line according to their type, and this is without loss of generality. Now suppose an agent i is considering to form a link with an agent that is distant two steps from i. Recalling best reply functions, $x_i^\star = \alpha\{\bar{x_j}\}_{j \in N(i)} + (1-\alpha)\theta_i$, and that $0 \leq \alpha < 1$ we have that in equilibrium action is chosen as a linear combination of type θ_i and average behaviour of the reference group. The restriction on the parameter α rules out situations in which actions overlaps in equilibrium, unless there are two agents i, j such that $\theta_i = \theta_j$. In such environment there are no benefits from possible deviations, and therefore we have that

$$
\begin{aligned}
i) \qquad & x_i > x_j \quad \Leftrightarrow \theta_i > \theta_j \\
ii) \qquad & x_i < x_j \quad \Leftrightarrow \theta_i < \theta_j \\
iii) \qquad & x_i = x_j \quad \Leftrightarrow \theta_i = \theta_j
\end{aligned}
\tag{B.1}
$$

To prove these claims, we could use a contradictory argument. Take for instance i), and assume that the inequality does not hold and therefore $x_j \geq x_i$. This can only be possible if $\{\bar{x}_h\}_{h \in N_j} > \{\bar{x}_{h'}\}_{h' \in N_i}$. Now take $j = \underline{\theta}$ and let i be the next player on the line of types and further assume that agent j is in a PS scenario. We then analyze the case in which

$$
\begin{aligned}
x_j &> x_i \\
x_j - x_i &> 0
\end{aligned}
\tag{B.2}
$$

and using equilibrium behaviors to rewrite

$$
\begin{aligned}
\alpha\{\bar{x}_h\}_{h \in N_j} + (1-\alpha)\underline{\theta} - \alpha\{\bar{x}_{h'}\}_{h' \in N_i} - (1-\alpha)\theta_i > 0 \\
\alpha(\{\bar{x}_h\}_{h \in N_j} - \{\bar{x}_{h'}\}_{h' \in N_i}) > (1-\alpha)(\theta_i - \underline{\theta})
\end{aligned}
\tag{B.3}
$$

Now we focus on $\{\bar{x}_h\}_{h \in N_j}$ and $\{\bar{x}_{h'}\}_{h' \in N_i}$. In particular $\{\bar{x}_h\}_{h \in N_j} = \frac{1}{k}\sum_{h=1}^{H} x_h$. The lowest possible value that $\{\bar{x}_{h'}\}_{h' \in N_i}$ can take is $\{\bar{x}_{h'}\}_{h' \in N_i} = \frac{1}{k}\sum_{h'=1}^{H'}$ with $H' = H$. Both agents have the same number of connections, but j is linked with i (implying i is linked with j). It is easy to note how any $H' > H$ would increase $\{\bar{x}_{h'}\}_{h' \in N_i}$, contradicting the fact that it is the lowest value that it can take. On the other hand any $H' < H$ would not be pairwise stable, given that we assumed j is in a PS scenario and that this would mean that the latter agents ignored agent i in the linking formation process. Nevertheless, deleting a link with agent j and adding a link with agent i would yield them a higher utility, which contradicts the pairwise equilibrium conditions. Therefore $H' = H$. Now getting back to the inequality

$$
\alpha(\{\bar{x}_h\}_{h \in N_j} - \{\bar{x}_{h'}\}_{h' \in N_i}) > (1-\alpha)(\theta_i - \underline{\theta})
\tag{B.4}
$$

we see that the left hand side simplifies to

$$
\frac{1}{k}(x_i - x_j)
$$

since the sets of links differ only by one connection. Given the initial assumption of $x_j \geq x_i$ is never positive. On the other hand the right hand side of B.4 is always positive, since $\theta_i > \underline{\theta}$ and $\alpha \in [0, 1)$. Therefore equation B.4 is impossible, unless $x_i > x_j$, which contradicts the initial assumption.

Using a similar argument we can show *ii)* and *iii)*, and we can iterate the reasoning to see that this applies to all $i \in N$. Moreover now it is easy to check that if an agent i finds beneficial to form a link with an agent distant two steps from his/her position (call it $i+2$), it must be, a fortiori, that i formed a link with the adjacent player (call it $i+1$). Indeed

$$U_i(g + i(i+2)) > U_i(g) \Rightarrow U_i(g + i(i+1)) > U_i(g)$$

The reverse is not necessarily true. It is also that for any $\theta_{i+2} > \theta_{i+1}$, everything else equal

$$U_i(g + i(i+1)) > U_i(g + i(i+2))$$

Moreover by symmetry

$$U_i(g + i(i+1)) > U_i(g) \Rightarrow U_{i+1}(g + (i+1)(i+2) > U_{i+1}(g)$$

.

Note that we have not used any additional assumption on δ, Θ and α with respect to those stated in the very beginning of the exposition of the model. this characterization is therefore independent by the values assumed by these variables. This concludes the proof of this Proposition. □

B.3 Proof of Remark 10

Proof. Consider first the case with $\alpha \to 1$. In such case we have that agent's behaviours are driven (almost) solely by social interactions. In equilibrium it will arise a unique action by all agents, and this can be easily checked studying best replies. Note that if α is even an ε smaller than 1, uniqueness of the Nash equilibrium is ensured. Precisely therefore there would not be a unique action, but the set of actions clearly shrinks toward a singleton for ε sufficiently small. Nevertheless, the cost of bearing links approaches 0, and therefore forming a link is beneficial for every $\delta > 0$. The complete network yields the maximum level of utility for every agent, and also the society as a whole experiences the highest level of welfare achievable. Clearly this network is pairwise stable, since no agent would delete any link, while connections are formed with all other agents.

Now we pass to the case with $\alpha \to 0$. It is rather the opposite case with respect to the previous one in terms of equilibrium behaviors, even though the resulting network is the same. However agents now do not care about conformism, i.e. they do not pay cost for deviations from other's behaviours. In equilibrium then every agent would respond to their identity (θ) and therefore $x_i = \theta_i$ for all i. Cost of forming links is again 0 and previous reasoning applies here as well.

Note that in both cases the result hold for every possible distribution of θs. □

B.4 Proof of Proposition 13

Proof. I prove the result by induction.

To proceed with the proof we have to show that there is a network g that is a pairwise farsightedly stable set. This could be if and only if for every $g' \in \mathbb{G} \setminus \{g\}$ we have that $g \in F(g')$, which denotes that network g belongs to an improving path departing from g'. Moreover it is uniquely pairwise farsightedly stable if and only if $G = \{g \in \mathbb{G} | F(g) = \emptyset\}$ and for every $g' \in \mathbb{G} \setminus G, F(g') \cap G \neq \emptyset$.[1] In other words we have to show that there is a unique pairwise farsightedly stable set, and it is composed by singletons.

It is useful in the first place to set the definition for single-peaked preferences in own links.

Definition 21. *We have single-peaked preferences in own links when for all $G'', G' \in \mathbb{G}$ such that G'' and G' are the set of networks that involve $|N(i; G'')| > |N(i; G')|$ or $|N(i; G'')| < |N(i; G')|$, and such that $G'' < G' \leq G^{B_i}$ or $G'' > G' > G^{B_i}$ we have that $U_i(G'', \theta_i) < U_i(G', \theta_i)$, for all $i \in S$, where G^{B_i} denotes the bliss point for all agents i in a coalition S.*

The definition states that agents have single-peaked preferences if they have a bliss point in preferences in own links, regardless of other agents choices.[2] Therefore this is a concept that involves set of networks, because we are able to fix only choices of links for agents with such bliss point. Note that the definition is reported allowing for coalitions, but in the present framework coalitions are composed by single agents. Coalition in the present context may arise allowing for more than one agent to be endowed with a same type.

[1] These are results found in Herings et al. (2009), in particular in Propositions 4 and 5, pag. 533.

[2] This is a novel concept in networks, although it is simply an adaptation of the previously known definition of single peaked preferences.

If there is a coalition with single-peaked preferences in own links, the set of networks G^{B_i} is a pairwise farsightedly stable set, meaning that all possible deviations from a network in $g \in G$ are deterred by a credible threat of ending worse off, and there exist a farsighted improving path from any network outside the set leading to some network inside the set. Presence of agents with single-peaked preferences is ensured by construction, because of peripheral players, endowed with extremes of the support of the distribution of θ. Such agents could have only asymmetric groups. Denoting such players with 1 and N, respectively for the agent in the left and the right of the distribution, we have that them would not deviate from their bliss point, and therefore this set constitutes an ending point. Indeed for any $g \notin G^{B_i}$ we have that $F(g) \cap G^{B_i} \neq \emptyset$. This is because for every network that involves the peripheral agents, it is possible to select those that belong to the set, because farsightedly they are the ones from which they are not willing to deviate. If otherwise there are networks where such peripheral players are not involved, they will be deleted from the network and other peripheral agents will be present.

Given the existence of such stable set of networks, we are able to refine more the configuration through equilibrium behaviours described in Proposition 9. In particular we have to show that the set is composed by a singleton, and therefore for every $g \in \mathbb{G} \setminus g^*$ we have $g^* \in F(g)$, meaning that there is a farsightedly improving path from every network $g \neq g^*$ that leads to g^*. We show that proving that g^* can be reached (1) from an empty network by addition of links and (2) from the complete network through deletion of links. Results follows from single-peakdness of preferences of peripheral agents.

(1) Consider the empty network. Agents add sequentially links towards a complete network. Sequentiality is due to Lemma 9, and therefore closest neighbours will be considered. The process ends for the peripheral players because for them would not be beneficial to add further links. Closest nodes to peripheral will always be willing to form links with peripheral because this would reduce the cost of forming further links. Iterating the reasoning we end up to g^* that is pairwise Nash stable. There exist coalitions for which it would be profitable to add further links, but always exist agents for which the resulting network configuration would not be profitable. By pairwise stability they cannot be reached.

(2) Now consider the complete network. Peripheral players are always willing to delete links, unless their bliss network is the complete network. In such case it is the only pairwise stable network. Given this deletion, the other agents will adjust their link configuration reaching again to a pairwise stable network g^*.

Through (1) and (2), we have that for all $g \neq g^*$, $g^* \in F(g)$. Therefore checking definition 12 it is easy to see that g^* satisfies all conditions. □

B.5 Proof of Proposition 14

Proof. In this proof I show how equilibrium behaviors are shaped when $\lim_{n_2\to\infty}$ while we let $n_1 < \infty$. Then I plug the obtained results into the utility functions, take the difference with utility after forming the link and before the link is formed, and find conditions on θ_2 such that this difference is positive, which means that forming new connections is profitable.

Trivially, we let $\lim_{n_2\to\infty} x_2 = x_2$, meaning that no matter how large is n_2, behaviors will not depend on that, and are always defined. However we know that $x_2^* = \alpha\{\bar{x}_j\}_{j\in N_2} + (1-\alpha)\theta_2$, with $\{\bar{x}_j\}_{j\in N_2} = \frac{1}{n_2}\sum_{j\in N_2} x_j = \frac{1}{n_2}\sum_{k=1}^{n_2-1} x_2 + \frac{1}{n_2}x_1$. Therefore we have that

$$\lim_{n_2\to\infty} x_2^* = \alpha x_2 + (1-\alpha)\theta_2 = \theta_2 \tag{B.5}$$

where we applied the Cesàro Mean Theorem to derive that $\frac{1}{n_2}\sum_{k=1}^{n_2-1} x_2 = x_2$. The argument follows directly from the fact that $\lim_{n_2\to\infty} x_2 = x_2$ as we stated in the beginning. Then $\frac{1}{n_2}x_1$ clearly vanishes, and it is how equation B.5 is found. Pasting B.5 into the utility function we get

$$\begin{aligned} U_2 &= \delta n_2 - \alpha\frac{1}{n_2}\sum_{j=1}^{n_2}(x_2 - x_j)^2 - (1-\alpha)(x_2-\theta_2)^2 \\ &= \delta n_2 - \alpha(\frac{1}{n_2}\sum_{j=1}^{n_2-1}(x_2-x_2)^2 + \frac{1}{n_2}(x_2-x_1)^2 - (1-\alpha)(\theta_2-\theta_2)^2 \\ &= \delta n_2 \end{aligned} \tag{B.6}$$

the second term on the right hand side vanishes because $x_2 = x_2$, the third vanishes since $\frac{1}{n_2} = 0$ at the limit, and the last vanishes since we used equilibrium behavior $x_2 = \theta_2$.

This result allows to simplify consistently the analysis, without losing the intuition behind the exercise. We just stated that for growing n_2 forming a link becomes less and less costly, until such cost vanishes entirely at the limit. Provided that we can now focus on the minority. Equilibrium behavior adding one link toward the majority would be

$$\begin{aligned} x_1^* &= \alpha\{\bar{x}_j\}_{j\in N_1} + (1-\alpha)\theta_1 \\ &= \alpha(\frac{n_1-1}{n_1}x_1 + \frac{1}{n_1}\theta_2) + (1-\alpha)\theta_1 \\ &= \frac{\alpha}{n_1 - \alpha(n_1-1)}\theta_2 \end{aligned} \tag{B.7}$$

where we used $\theta_1 = 0$ to simplify the algebra. As before we paste now into utility function af agents 1.

$$\begin{aligned} U_1 &= \delta n_1 - \alpha \frac{1}{n_1}(\frac{\alpha}{n_1 - \alpha(n_1-1)}\theta_2 - \theta_2)^2 - (1-\alpha)(\frac{\alpha}{n_1 - \alpha(n_1-1)}\theta_2)^2 = \\ &= \delta n_1 - \frac{\alpha}{n_1}(1 - \frac{\alpha}{n_1 - \alpha(n_1-1)})^2\theta_2^2 - (1-\alpha)(\frac{\alpha}{n_1 - \alpha(n_1-1)})^2\theta_2^2 \end{aligned} \tag{B.8}$$

Now, we know that with no addition of the further link $U_1' = \delta(n_1 - 1)$, since agents are linked with $n_1 - 1$ other agents, and links are costless. Therefore we take now the difference between $U_1 - U_1'$, we want that this is positive, meaning that the addition of the link is profitable, and then we solve for θ_2. As a result we obtain

$$\theta_2 < \sqrt{\frac{\delta}{(\frac{\alpha}{n_1}(1 - \frac{\alpha}{n_1 - \alpha(n_1-1)})^2) + (1-\alpha)(\frac{\alpha}{n_1 - \alpha(n_1-1)})^2}} := \xi(\delta, \alpha, n_1) \tag{B.9}$$

□

B.6 Proof of Remark 19

Proof. To show this result, we have to solve the recursive formula.

$$\begin{aligned} \mathbf{x} &= \alpha\mathbf{W}\mathbf{x} + (1-\alpha)\mathbf{t} \\ \mathbf{x} &= \alpha\mathbf{W}(\alpha\mathbf{W}\mathbf{x} + (1-\alpha)\mathbf{t}) + (1-\alpha)\mathbf{t} \\ \mathbf{x} &= \alpha^2\mathbf{W}^2(\alpha\mathbf{W}\mathbf{x} + (1-\alpha)\mathbf{t}) + \alpha\mathbf{W}(1-\alpha)\mathbf{t} + (1-\alpha)\mathbf{t} \\ &. \\ &. \\ \mathbf{x} &= \sum_{k\geq 0}^{\infty}\alpha^k\mathbf{W}^k + \sum_{k\geq 0}^{\infty}\alpha^k\mathbf{W}^k(1-\alpha)\mathbf{t} \\ \mathbf{x} &= (\mathbf{I} - \alpha\mathbf{W})^{-1}(\mathbf{I} + (1-\alpha)\mathbf{t}) \end{aligned} \tag{B.10}$$

□

B.7 Example - Three agents

This particularly simple example is meant to provide an example of how farsightedness helps in avoiding cycles into the framework I have described. When only $n = 3$ agents are in the network, the whole possible structures are given by $2^3 = 8$. However, thanks to

Proposition 9 we can reduce the number of feasible networks, because I have shown that only certain structures can arise in equilibrium. Figure B.1 shows that.

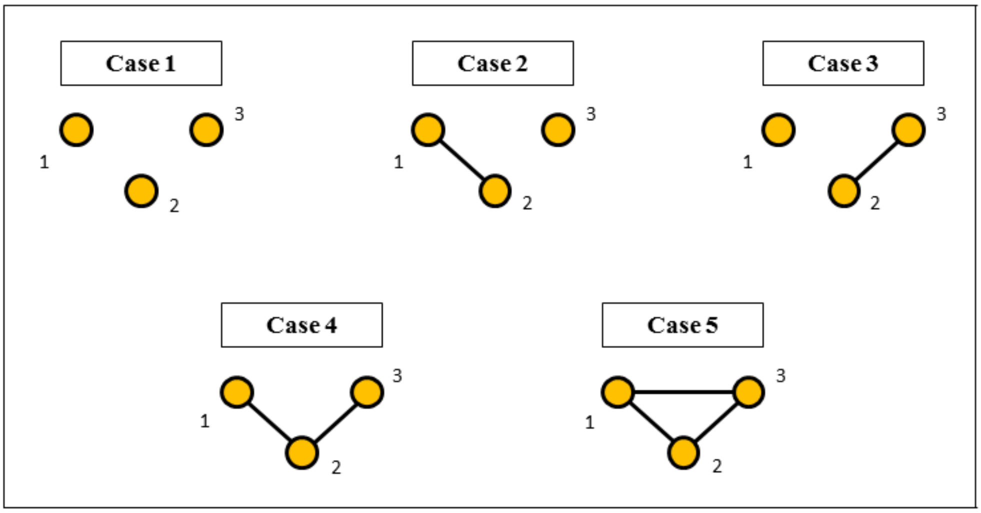


Figure B.1: The figure represents all the possible network configurations when 3 agents are considered. Some networks are excluded from the figure because of considerations in Proposition 9.

Provided that, this simple environment allows us to present results under a fairly general framework. For bigger n we would have too many networks to compare. With such restriction on population size, instead, we can do comparative statics on all parameters of the model.

Basic conditions for agents to start to form connections are that agents at the periphery are at least willing to form link with the middle agent. In next Remark these conditions are formalized.

Remark 22. *With $n = 3$ agents,it is profitable for both involved agents to form a link when the following are true*

- *Agent 1 and 2 are connected (g_2) whenever* $(\theta_1 - \theta_2)^2 < \frac{(1+\alpha)^2}{\alpha-\alpha^2}\delta$.
- *Agent 2 and 3 are connected (g_3) whenever* $(\theta_2 - \theta_3)^2 < \frac{(1+\alpha)^2}{\alpha-\alpha^2}\delta$

The results here are intuitive. If we have a population of 2 agents, we have that they will form a link whenever the benefits outweigh the costs, which are determined through the distance of these agents in terms of their characteristics. Moreover, the benefit (or the perception of it) depends on α. In particular for mid ranges of the parameter (between 0.2 and 0.5) the benefit keeps relatively low, meaning that social interactions are chosen more carefully by agents. Whenever out of this range instead the benefits explode, meaning that agents do not care about other's characteristics, both because they would just conform to them, either because they would just respond to their types.

Now we explore whether it is profitable for agent 2 to add a connection on top of one of the two scenarios above depicted. This is the case where farsightedness hinges in the model. To see this, assume we are in network g_2 where 1 and 2 are connected, but agent 3 is left apart. However imagine that 2 and 3 would both benefit form adding a connection. By pairwise stability they would add this link ending up in g_4. However now two different situations may arise: *(1)* agent 1 has still a positive utility, and therefore severing a link would not be optimal, since he would end up alone, with 0 utility. In this case the network is PS. *(2)* Agent 1 is now facing negative utility, because the indirect influence he receives from agent 3 pull him too far from her own identity. In such a case agent 1 would be better off severing the link, and the network configuration would then be g_3.

From now on agent 2 has again 2 options: *(2_a)* agent 2 is better off in g_3 than in g_2. In such a case g_3 is PS. *(2_b)* Agent 2 is better off in g_2 with respect to g_3. Now we would have a cycle, and this is the only case where farsightedness applies. In fact rather than cycling on g_2, g_3, g_4, agent 2 would anticipate that and will not form the connection with agent 3 in the first place, letting g_2 be FPS.

To simplify exposition of further results we let $\alpha = 0.5$, $\delta = 1$ and $\theta_1 = 0$.[3]. All results are presented in the following table which summarizes all conditions for every single network to be stable, under the above mentioned restrictions on parameters.

B.8 Alternative utility function

As a final exercise I want to discuss an alternative utility function. It is reasonable to contest the specificity of the parametric form of the utility function. However, what really characterize the game here is the best reply function. We can get the same best reply changing slightly the utility function used in Calvó-Armengol et al. (2009). In particular in the mentioned paper, agents' own preferences are modeled disjointly from the action that may derive from peer's interactions, and that was made for measurement purposes, when approaching the data. Here instead the two sources are measured on the same metric. Theoretically, in fact, there should be no difference between the action one would exert in isolation, and the component of behavior one undertakes because of social

[3]This are strong restriction on parameters, but algebra becomes extremely tedious letting all parameters free to move. Nevertheless, $\delta = 1$ and $\theta_=0$ are without loss of generality, and $\alpha = 0.5$ is in my opinion the most interesting case. This can be partly motivated by results in Remark 10

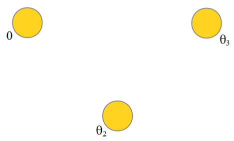
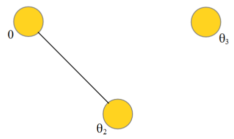
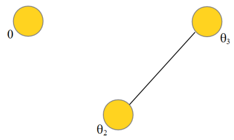
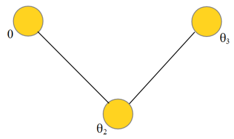
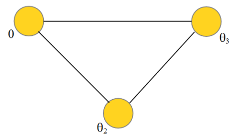
Conditions with $\alpha = 0.5$, $\delta = 1$ and $\theta_1 = 0$	
Networks	**Conditions**
	$\theta_2 > 3, \quad \theta_3 > 3\theta_2$
	$\theta_2 \in (0, 2.4], \quad \theta_3 > 12 - 4\theta_2$ $\theta_2 \in (2.4, 3), \quad \theta_3 > \theta_2$
	$\theta_2 \in (1.8, 2.4), \quad \theta_3 \in (12 - 4\theta_2, 3\theta_2)$ $\theta_2 > 2.4, \quad \theta_3 < 3\theta_2$
	$\theta_2 \in (0, 2], \quad \theta_3 < \frac{1}{5}(12 + 4\theta_2)$ $\theta_2 \in (2, 2.4), \quad \theta_3 < 12 - 4\theta_2$
	$\theta_2 = 2, \quad \theta_3 < 4.46$ $\theta_2 = 3, \quad \theta_3 < 4$ $\theta_2 = 3.5, \quad \theta_3 < 3.57$

Table B.1: This table shows conditions on θ_2, θ_3 under which every single network is PS (or FPS) assuming $\alpha = 0.5, \delta = 1, \theta_1 = 0$. Conditions on the same row must hold jointly, and were obtained deriving conditions for utility of all players to be greater than in any other scenario. The first row show the empty network denoted in the text with g_1. g_2, g_3, g_4, g_5 follow.

influence. However that paper provides mainly an empirical contribution and therefore making such distinction is important for the specification of the underlying econometric model.[4] Consider the following utility function

$$U_i = \sum_{j \in N_i(g)} \delta - \frac{1}{2}x_i^2 + \alpha \frac{1}{d_i(g)} \sum_{j \in N_i(g)} x_i x_j + (1-\alpha) x_i \theta_i \tag{B.11}$$

This utility function just states that utility is concave in own action, there is a complementarity between own action and own links' behavior, and a complementarity between own behavior and own type. Since types are exogenously given, it represents again the bliss point of an agent in undertaking a specific action. The interpretation of this parametric utility function is slightly different from the one proposed above. Here there are no more deviations from reference points, but complementarities are more evident. Here every agent faces the same disutility from action x, but the heterogeneity makes anyone of them to prefer a specific level of it. Best reply function are identical to those proposed before[5] and therefore given by

$$x_i^\star = \alpha \{\bar{x}_j\}_{j \in N_i(g)} + (1-\alpha)\theta_i \tag{B.12}$$

Even though the model is similar in its characteristics with the one presented above, the interpretation is definitely different. Homophily still arises, but the underlying motives are different. Imagine a high type, i.e. an agent with high θ already connected with similar agents. Connecting with other people would push his choice of x under his/her own reference point. This would generate several effects. The disutility from action would go down, increasing the agent utility. The term that captures complementarity effects with other peers would go down because the average contribution of peers to "team work" would be lower. Finally also the idiosyncratic term would go down because the agent exerts now a lower level of x. Summing up all these effects, the net value may be negative. Therefore connections would be beneficial if and only if the benefit of such further links is high enough to compensate for the loss.

Summing up connections with similar agents will always be preferred to those with agents with distant characteristics. In fact, given the same benefit deriving from the connections, the net utility loss deriving from the other terms would be lower. Again,

[4]One may take as benchmark the model in Ballester et al. (2006) and note as a difference the simple introduction of an exogenous heterogeneity in the identity of players.

[5]See equation 2.3.

depending on the value function that describes the benefit from connections, there are links that is not profitable to make, and the more an agents' characteristics are far from another agent, the more probable would be that a connection takes place between them.

Appendix C

Search over the network

C.1 Expected Degree

$$E(d_{i1}) = E_i(p_{ij}) + E_i(p_{ij}) \sum_{k \in N \setminus i,j} \frac{p_{ki}}{N-1} + (1 - E_i(p_{ij})) \sum_{j \in N \setminus i} \frac{p_{ji}}{N-1} =$$
$$= E_i(p_{ij}) + E_i(p_{ij}) \sum_{k \in N \setminus i,j} \frac{p_{ki}}{N-1} + (1 - E_i(p_{ij}))E_i(p_{ji})$$

For large N, or for N way larger than T, we can approximate $\sum_{k \in N \setminus i,j} \frac{p_{ki}}{N-1} \approx E_i(p_{ji})$, such that 3.5 can be rewritten as

$$E(d_{i1}) = E_i(p_{ij}) + E_i(p_{ji}) = 2E_i(p_{ij})$$

where the last holds under undirected network. Iterating the formula we can calculate the expected degree in any period $t+1$, which depends on the expected degree in period t, since this is a cumulative process. Let us call $\bar{p}_i = E_i(p_{ij}) = \frac{1}{N-1} \sum_{j \in N \setminus i} p_{ij}$

$$E(d_{it+1}) = E(d_{it}) + \bar{p}_i \tag{C.1}$$

Therefore the general expected degree at time t is given by

$$E(d_{it}) = 0 + 2t\bar{p}_i = 2t\bar{p}_i \tag{C.2}$$

C.2 Stopping time

$$U_{it} \leq EU_{it+1}$$

thus search continues whenever

$$(T-t)d_{it} \leq (T-t-1)(d_{it}+2\bar{p}_i)$$

after some algebra we get

$$d_{it} \leq (T-t-1)2\bar{p}_i$$

from which we can get

$$E(t^*|d_{it}) = T - \frac{d_{it}}{2\bar{p}_i} - 1 \tag{C.3}$$

and

$$\bar{d}_{i|t} = 2(T-t-1)\bar{p}_i \tag{C.4}$$