CENTRO DE INVESTIGACIÓN Y DOCENCIA ECONÓMICAS, A.C.

CIDE

THE EMERGENCE OF NETWORKS IN THE PRESENCE OF HOMOPHILOUS AND HETEROPHILOUS PREFERENCES

TESINA

QUE PARA OBTENER EL GRADO DE

MAESTRO EN ECONOMÍA

PRESENTA

ISRAEL COLCHADO FLORES

DIRECTORA DE LA TESINA: DRA. ISABEL MELGUIZO LÓPEZ

CIUDAD DE MÉXICO
2020

To my family. Your eternal support, love and teaching made possible this achievement.

Aknowledgements

Abstract

Human beings make links with similar and dissimilar people based on certain attributes. In this thesis, I propose a network formation model where agents have a dichotomic attribute, they have preferences for similar and dissimilar individuals, and there is also a linking cost. Three different networks are characterized using the concept of pairwise stability. As a result, agents link with similar or dissimilar individuals if and only if their preferences reflect this behavior. Finally, a discussion about the limitations and future research of this model is made.

Key words: Homophily, heterophily, social networks, friendship, pairwise stability.
JEL classification: A14, Z13

Contents

List of Figures

Chapter 1

Introduction

As social animals, human beings need to interact with other people to guarantee safety, companionship, efficiency in their daily life tasks, etc. Evidence suggests that individual characteristics determine tie formation.

Homophily is the tendency of individuals to interact with similar people. McPherson et al. (2001) define it as "the principle that contact between similar people occurs at a higher rate than among dissimilar people". In Lazarsfeld et al. (1954), the famous quote "birds of the feather flock together" summarizes the empirical effects of homophily. For instance, white American students often make friends with other white schoolmates.[1]

Even when there is much evidence of homophily in the scientific literature, it is also true that there are several situations in which it is not the dominant tendency. Heterophily is a sociological concept whose main characteristic is the existence and maintenance of relations among dissimilar individuals. See Lozares et al. (2014), Barranco et al. (2019). Many of these interactions are based on complementarities among different skills/types and they impact positively on

---

[1]McPherson et al. (2001) remark that homophily in race and ethnicity determines strongly how we interact in our daily surroundings. Other attributes are age, religion, education occupation and gender in this order. Examples of homophilous tendencies can be found in friendship Currarini et al. (2009), Lazarsfeld et al. (1954), marriage Kalmijn (1998) and certain labor relations Ibarra (1992).

communication through weak ties Granovetter (1977).[2,3]

Sometimes, some individuals are homophilous and other individuals are heterophilous. For instance, some Britpop musicians prefer to make professional and sentimental relationships with other musicians from their same origin social class, whereas other musicians prefer to engage with dissimilar individuals Millward et al. (2017). Another example is found in the ties between customers and frontline employees, like waiters or clerks Streukens and Andreassen (2013). Introvert customers prefer to interact with extroverted employees and more extroverted customers have stronger preferences for extroverted employees.

People take advantage of homophilous and heterophilous links in real life. For instance, when unemployed individuals are seeking a job, the heterophilous interactions beyond their strongest ties play a fundamental role in getting a new job. See Lin et al. (1981), Granovetter (1970). However, evaluators tend to rank highly candidates similar to them unless they are well informed about the extent of intergroup differences and they may condition their assessment in candidates' group belonging Bagues and Perez-Villadoniga (2013).

Homohpily and heterophily also play an important role in effective communication. Receivers and senders are homophilous on some relevant attributes because people tend to trust in similar individuals. But they cannot be so similar because if they both were completely homophilous, such interaction would be redundant Rogers and Bhowmik (1970). For instance, the interaction between women of different social classes increases awareness of different family planning methods in low class women Liu and Duff (1972).

The objective of this thesis is to understand how homophilous and heterophilous preferences determine network formation, for example in friendship relationships. Particularly, this work focuses on the phenomenon found in Patacchini and Zenou (2016). In this paper, Patacchini and Zenou analyze data about friendship networks among high-school students in black integrated

---

[2]Examples of heterophily can be found in investment banks' tie formation Podolny (1993), interdisciplinary teams like academic collaborations Jha and Welch (2010), Jones et al. (2008), Moody (2004), Newman (2000) and effective risk-sharing among farmers Bramoullé and Kranton (2007), Fafchamps and Gubert (2007).

[3]Other examples related to effective communication are Korte and Milgram (1970) and Kerckhoff and Back (1965), where weak ties reduced the time in diffusion processes.

schools in America using the Coleman's homophily index. This index is a normalized measure that quantifies how homophilous are the links of each individual in a network, whose values are bounded between 1 and -1. The more positive it is, the more the individual shows a homophilous trend in his interactions. The opposite is also true, a more negative measure implies a dominance of heterophilous links. Scientists found that white students have strong homophilous interactions as it can be seen in the right side of figure 1, in which the most frequent Coleman's homophily index is positive and near form one for white students. However, the distribution of the Coleman's index is bimodal for black students. This means that some black students interact more frequently with black friends and some others make friendship links with white students, as it can be seen in the left side of figure 1.
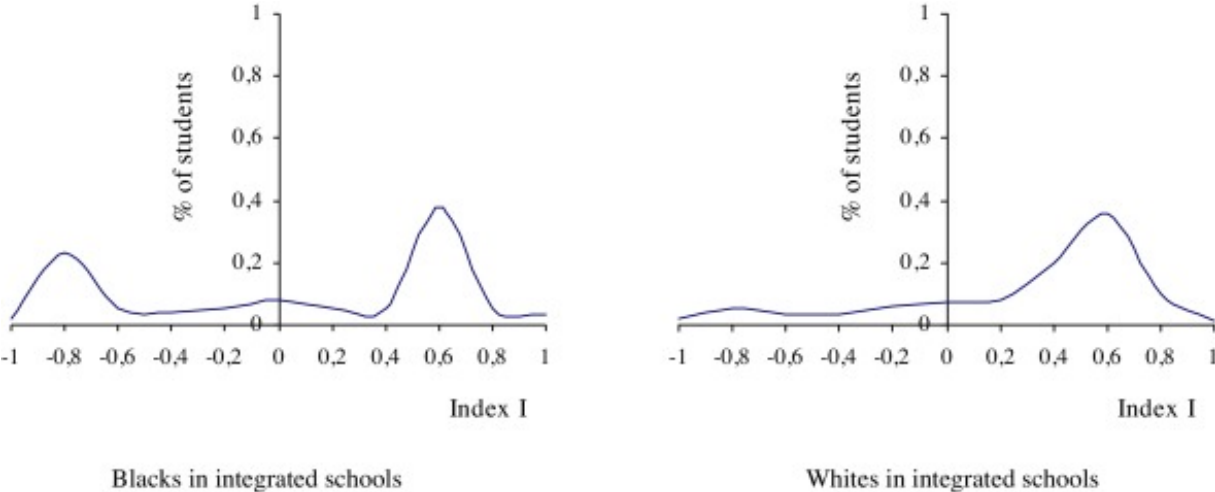
Figure 1: Distribution of black (left) and white (right) students according to their Coleman indexes. Source: This figure is taken from Figure 1 in De Marti and Zenou (2017).

In De Marti and Zenou (2017), the authors propose a network formation model "that can explain the socialization patterns in figure 1". However, the networks they characterize do not show this bimodality in the Coleman's index, except for particular examples (that will be exemplified in the following section).

Following the spirit of De Marti and Zenou (2017), I propose a network formation model that reproduces this bimodality. In this model, agents have a dichotomic attribute, splitting individuals into two disjoint groups. It is supposed that individuals like to have similar friends

as well as dissimilar friends. The model allows that the strength of these two opposite forces can be individual specific, i.e there is *ex ante* heterogeneity in individuals' preferences. There is also a constant linking cost.

Regarding network formation, it is considered the emergence of pairwise stable structures as in the influential approach of Jackson and Wolinsky (1996). Intuitively, a pairwise stable network is one in which connected individuals do not want to break their links and for any two not connected individuals, at least one of them does not want to be linked with the other. I will also calculate the Coleman's index when this equilibrium is reached.

The main results characterize several pairwise stable networks. A network in which individuals link with all the agents is pairwise stable if and only if individuals have not extreme homophilous and heterophilous preferences, and the linking cost is not so high. It is also proved that the only pairwise stable network in which individuals only link with similar agents is the network where individuals link in dyads. This segregated network is pairwise stable if and only if the individuals' preferences are sufficiently homophilous and the linking cost is relatively high. Finally, I characterize a network that exhibits bimodality in Coleman indexes for one group and unimodality for the other groups, characterizing the leading example in De Marti and Zenou (2017).

For future research, it will be of interest to make some comparative statics, analyzing the transition between theses pairwise stable networks as the homophilous and heterophilous preferences varies. It would also be interesting to propose a new model where homophilous and heterophilous preferences emerge due to an underlying process that do not explicitly assumes homophily. For instance, a model could assume that individuals interpret better information coming from similar people. As a consequence, individuals could make homophilous links.

Chapter 2

Other Related Literature in Economics

The proposed model is inspired by several models in the scientific literature. Dasgupta and Goyal (2009) propose a model in which individuals have two attributes. They have to choose whether they join into two groups, only one group or none of them. In contrast to the proposed model in this thesis, the decision that individuals pursue in their model is the adherence to groups and not individual linking decisions. Moreover, they use the Nash equilibrium concept whereas the pairwise stability equilibrium concept is used in this thesis. The focus is also different, because I look at the opposing forces of homophily and heterophily in network formation, and Dasgupta and Goyal look at the emergence of groups based on the further division of the prospect generated by this group. An example is a group that produces a public good. Then, every individual benefits from the public good as whole.

This work has a similar approach to the model proposed in Iijima and Kamada (2017). They also propose a network formation model and focus on the formation of pairwise stable networks. In this model, agents have multiple continuous attributes, their utilities depend on the distance between direct and indirect connected individuals and the costs to be directly linked with other agents. However, the authors model a situation in which individuals with multiple attributes benefit less the higher is the social distance, defining a measure of how different attributes are between two agents. Also, their main focus is on the macro properties of stable networks, as

clustering and average path length, whereas this thesis focuses on the interactions of agents in a micro level.

The model proposed in this work assumes that individuals care about the ratio of similar and dissimilar friends. This choice is based on the proposal by Schelling (1971). In Schelling's proposal, there are two types of agents where each agent belongs to a certain group. Individuals are happier in their neighborhood when the ratio of similar agents is above a certain threshold than when this ratio is below it. If an individual is unhappy in his neighborhood, he moves until he belongs to a neighborhood where his ratio is above this threshold. The main result is that segregation emerges even when individuals have these mild preferences. Moreover, a minority group tends to become more segregated from the majority group when its relative size diminishes. The main difference with the proposed model is the notion of network, which is crucial in this model whereas there is no notion of network in the Schelling's model.

Finally, the proposed model is closely related to the model proposed by Currarini et al. (2009). As in their model, it is assumed preferences for similarities and differences. Moreover, each individual only recognizes if his friends are similar or different from his type (dichotomic attributes). This last assumption is based on empirical evidence, because the main form of differentiation among individuals is based on similar-different assessments, and no other elaborated forms of stratification. See McPherson et al. (2001) and Marsden (1988). However, the Currarini and Jackson's model is one of dynamic matching based model and ours is one shot network formation model.

Chapter 3

Model

3.1 Set up of the model

Let $N = \{1, 2, .., n\}$ be the finite set of individuals. Every individuals has a dichotomic attribute which takes values A or B. For that reason, there are two types in this model because individuals split into two disjoint groups according to their attributes. This means that for every individual $i \in N$, this individual belongs to N_k for some $k \in \{A, B\}$, where the set N_A represents those individuals with attribute A and the set N_B represents those individuals with attribute B. Notice that $N_A \cup N_B = N$. Let n_A and n_B their respective cardinalities and without loss of generality $n_A \geq n_B \geq 2$. Finally, it is defined $N_{-k} = N - N_k$ for all $k \in \{A, B\}$.

The set of all possible networks is denoted as g^N. The subset of N which only has elements i and j is represented as ij. This notation also represents an undirected link between individuals i and j, that is $ij = ji$. As usual, $ij \in g$ means that i and j are connected under the network g and $ij \notin g$ means that i and j are not connected under the network g. The network $g + ij$ is the network g to which it is added the link ij and the network $g - ij$ is the network g to which it is removed the link ij.

Taking a fixed network $g \in g^N$, for every $i \in N$ it is defined the set of his friends as $N_i(g)$, the set of his similar friends as $S_i(g)$ and the set of his dissimilar friends as $D_i(g)$. Formally:

- $N_i(g) = \{j \in N \mid ij \in g\}$

- $S_i(g) = \{j \in N \mid ij \in g, \; i, j \in N_k \text{ for some } k \in \{A, B\}\}$

- $D_i(g) = \{j \in N \mid ij \in g, \; i \in N_k, \; j \in N_{-k} \text{ for some } k \in \{A, B\}\}$

Their cardinalities are defined as $n_i(g) = |N_i(g)|$, $s_i(g) = |S_i(g)|$ and $d_i(g) = |D_i(g)|$. When there is no confusion on network g, there will be expressed their cardinalities simply by n_i, s_i and d_i respectively.

All individuals like linking with similar and dissimilar agents. Individuals are potentially heterogeneous in these preferences. Then, it is supposed an assumption close to Schelling's proposal Schelling (1971), where individuals only care about the ratios. They obtain more utility when these ratios increase, with decreasing marginal utility. These preferences are represented using the function $f : \mathbb{R}_+ \longrightarrow \mathbb{R}$, such as $f(x)$ increases with x and $f(x + 1) - f(x)$ decreases with x. It is supposed that $f(0) = 0$ and the linking cost is $c > 0$ for every agent.[1]

Formally, for every individual $i \in N$ there is a utility function $u_i : \{g \mid g \subset g^N\} \to \mathbb{R}$ such as:

$$u_i(g) = \lambda_i \, f(r_s^i(g)) + (1 - \lambda_i) \, f(r_d^i(g)) - c \, n_i(g)$$

where $r_s^i(g) = \frac{s_i(g)}{n_i(g)}$ is the ratio of similar friends of i in the network g, $r_d^i(g) = \frac{d_i(g)}{n_i(g)}$ is the ratio of dissimilar friends of i in the network g. Let the intensity of homophily as $\lambda_i \in (0, 1)$, for every $i \in N$. It is possible to assume a common parameter λ for all individuals and many of the results in this this will remain unaltered. However, it is necessary some heterogeneity in the intensities to characterize a network which exhibits a bimodality in distribution of the Coleman's indexes of some type of individuals.

Notice that the parameter $r_s^i(g)$ is defined as a homophily index in Currarini et al. (2009) and the intensity of homophily λ_i is a kind of homophily index in the preferences for each individual, a same type bias in the individuals' preferences.[2]

---

[1]This means that the cost function is linear, thus the marginal cost is constant.

[2]Currarini et al. define a same type bias in the preferences as those preferences whose return to additional

I use the definition of pairwise stable network. This definition was introduced by Jackson and Wolinsky (1996). In this thesis, it is defined a pairwise stable network as:

Definition 1. *A network g is pairwise stable if two conditions are satisfied:*

1. *For every pair of connected individuals in the network, both of them are better if they keep being connected than without their link. Formally,*

$$\forall ij \in g, \; u_i(g) \geq u_i(g - ij) \quad and \quad u_j(g) \geq u_j(g - ij).$$

2. *For every pair of non-connected individuals, at least one of them does not want to make a tie with the other. Formally,*

$$\forall ij \notin g, \quad if \quad u_i(g + ij) > u_i(g) \quad then \quad u_j(g + ij) < u_j(g).$$

In this thesis, it is used the definition 1 to characterize some networks. It is also defined some networks that will be analyzed.

Definition 2. *A network g is completely connected network if every individual is linked with all other agents. Formally, $N_i(g) = N - \{i\} \; \forall i \in N$.*

Definition 3. *A network g is fully intraconnected if all individuals of the same type are connected.*

Definition 4. *A network g is segregated if agents only link with similar individuals.*

---

friendship is larger when they have more similar friends. This definition can be adapted to the proposed model and say that an individual has same type bias in his preferences when his intensity of homophily is greater than a half.

3.2 Characterization of some equilibria

Let me clarify that the proofs of each proposition can be found in the Appendix A. Before enunciating the propositions, it will be introduced some important notation to characterize pairwise stable networks. Because there are a finite number of individuals, it is defined the greatest and the lowest λ_i values of sets N_A and N_B as:

$$\bar{\lambda}^k = \max_{i \in N_k} \lambda_i \quad \forall k \in \{A, B\}$$

$$\underline{\lambda}^k = \min_{i_k} \lambda_i \quad \forall k \in \{A, B\}$$

These notations will be used to establish upper and lower bounds of the intensity of homophily *per type*. The first result states that a fully intraconnected network cannot be segregated.

Proposition 1. *If a fully intraconnected network is pairwise stable, then it cannot be segregated.*

Proposition 1 means that for any fully intraconnected network, there must be some individuals who are linked with dissimilar people. This result accords with the analyzed data in Patacchini and Zenou (2016), where there are black and white students in integrated colleges who link with dissimilar classmates.

The previous result states that a network cannot be fully intraconnected and segregated at the same time. The next result sets that the only segregated pairwise stable network emerges when all individuals link in pairs or dyads.[3] A similar result was found in the co-author model in Jackson and Wolinsky (1996), where the network is strongly efficient if individuals are linked in dyads. For that reason, It will be assumed that n_A or n_B are even numbers.[4] The concept crossed link will be used to indicate a link between different types.

---

[3]They link in dyads because the utility function considers ratios. When individuals only link with similar or dissimilar individuals, adding new links does not change these ratios at all.

[4]When n_A and n_B are odd numbers, It is possible to find similar conditions in which all individuals link in dyads and those individuals who are not linked, they both prefer to be alone and do not make links with dissimilar individuals.

Proposition 2. *Let n_B even. The only segregated pairwise stable network is that in which individuals link in dyads (as the one in panel 2b). The network in dyads is pairwise stable if and only if the linking cost is intermediate and preferences are sufficiently homophilous. Thus, no individual breaks his same type link and even the more heterophilous individuals do not want to make a crossed link.*[5]

Formally, the expression "the linking cost is intermediate" means:

$$\max_{i \in N_k}\{0,\ f(1/2) - \lambda_i f(1)\} < c \leq \min_{i \in N}\{\lambda_i f(1)\} \qquad \textit{for some} \quad k \in \{A, B\},$$

and the expression "preferences are sufficiently homophilous" means:

$$\underline{\lambda}_k \geq \frac{f(1/2)}{2f(1)} \quad \textit{for the same} \quad k \in \{A, B\}.$$

Assertive mating can induce segregated networks in dyads. This way of mating means that animals and humans tend to choose similar sexual partners in phenotypes, because this increases the altruism among family members Thiessen and Gregg (1980), Thiessen et al. (1997). Moreover, human couples that had been together longer have more similar perceived personality traits Little et al. (2006).

Both results are represented in figure 2. No network has crossed links. However, the network in panel 2a is fully intraconnected, a result that is impossible according to proposition 1. The only segregated network possible in this model is formed by dyads, represented in panel 2b. Notice that this network is not fully intraconnected.

---

[5]It is possible that all individuals have the same homophilous preferences. In that case, all individuals have homophilous preferences and no one wants to make a crossed link.
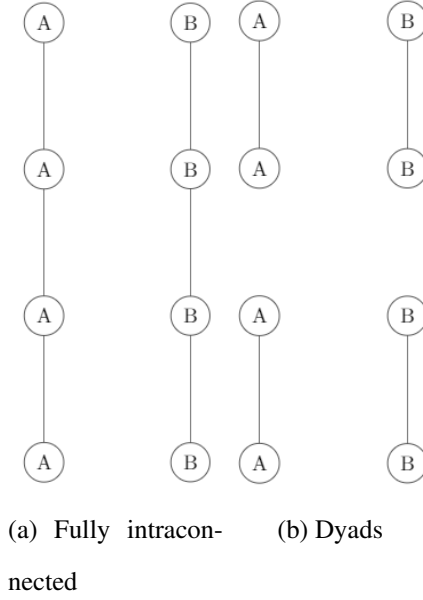
(a) Fully intracon-
nected

(b) Dyads

Figure 2: Segregated networks. Source: Own elaboration.

Before enunciating other propositions, it is introduced the next notation. For a network g, individual i has $n_i(g)$ friends, where $s_i(g)$ friends share his same attribute. Then, the marginal change in benefits of breaking a link with a similar individual is defined as:

$$\Delta_i^k s = f\left(\frac{s_i(g)}{n_i(g)}\right) - f\left(\frac{s_i(g) - 1}{n_i(g) - 1}\right) \geq 0 \qquad \forall i \in N_k, \quad k \in \{A, B\}.$$

This expression represents the marginal benefit that individual i gets in his homophilous preferences when he breaks with a similar individual. Notice that $\Delta_i^k s$ is decreasing in the ratio $r_s^i(g)$ because it is assumed that $f(x+1) - f(x)$ decreases with $x \in \mathbb{R}_+$. Intuitively, individuals get larger marginal benefits adding links with similar people when they have few similar friends than when they have many similar friends.

Now, imagine that we have a hypothetical situation where individual i makes a new link with a similar agent. This individual would have $n_i(g)+1$ friends and $s_i(g)+1$ similar friends. Then, the marginal change in benefits of adding a new link with a similar individual is defined as:

$$\Delta_i^k s' = f\left(\frac{s_i(g) + 1}{n_i(g) + 1}\right) - f\left(\frac{s_i(g)}{n_i(g)}\right) \geq 0 \qquad \forall i \in N_k, \quad k \in \{A, B\}.$$

It is important to distinguish between the expressions $\Delta_i^k s$ and $\Delta_i^k s'$. The former represents the real marginal benefit that individual i gets because it compares the homophilous component of his utility function taking as reference g with this same component taking as reference the network g when a link with a similar individual is broken. The later expression represents a hypothetical marginal benefit because it compares the homophilous component of his utility function when it is added a link with a similar individual to g with this same component taking as reference the original network g.

However, breaking or adding links with a similar individual also include a loss in his associated utility to his heterophilous component.

When agent i has $n_i(g)$ friends, the loss in his utility due to breaking a link with a similar individual is:

$$\tilde{\Delta}_i^k d = f\left(\frac{d_i(g)}{n_i(g)}\right) - f\left(\frac{d_i(g)}{n_i(g) - 1}\right) \leq 0 \qquad \forall i \in N_k, \quad k \in \{A, B\}.$$

This loss is related to the heterophilous part of the individuals's preferences. When individuals breaks his link with a similar individual, the ratio of dissimilar friends is bigger than the original ratio when this link is included. For that reason, individual i has a loss in his heterophilous preferences.

In the hypothetical case where individual i has $n_i(g) + 1$ friends, the loss associated to adding a new link with a similar individual is:

$$\tilde{\Delta}_i^k d' = f\left(\frac{d_i(g)}{n_i(g) + 1}\right) - f\left(\frac{d_i(g)}{n_i(g)}\right) \leq 0 \qquad \forall i \in N_k, \quad k \in \{A, B\}.$$

Similarly, for every agent $i \in N_k$, $k \in \{A, B\}$, the marginal change in benefits of breaking with a dissimilar individual when he has $n_i(g)$ friends and $d_i(g)$ dissimilar friends is defined as:

$$\Delta_i^k d = f\left(\frac{d_i(g)}{n_i(g)}\right) - f\left(\frac{d_i(g) - 1}{n_i(g) - 1}\right) \geq 0 \qquad \forall i \in N_k, \quad k \in \{A, B\}.$$

For the hypothetical case where individual i adds a new link with a dissimilar individual, the marginal change in benefits of adding a new link with a dissimilar individual is defined as:

$$\Delta_i^k d' = f\left(\frac{d_i(g) + 1}{n_i(g) + 1}\right) - f\left(\frac{d_i(g)}{n_i(g)}\right) \geq 0 \qquad \forall i \in N_k, \quad k \in \{A, B\}.$$

The loss in his utility associated to breaking or adding a link with a dissimilar individual are:

$$\tilde{\Delta}_i^k s = f\left(\frac{s_i(g)}{n_i(g)}\right) - f\left(\frac{s_i(g)}{n_i(g) - 1}\right) \leq 0 \qquad \forall i \in N_k, \quad k \in \{A, B\},$$

$$\tilde{\Delta}_i^k s' = f\left(\frac{s_i(g)}{n_i(g) + 1}\right) - f\left(\frac{s_i(g)}{n_i(g)}\right) \leq 0 \qquad \forall i \in N_k, \quad k \in \{A, B\}.$$

Finally, it is defined $n_{-k} \equiv n_{N-N_k}$ for k taking values A or B, this implies that $n_{-A} = n_B$ and $n_{-B} = n_A$.

Once it has been remarked that a segregated and fully intraconnected network is not possible in this model, I characterize a particular case of full intraconnection: when the network is completely connected.

Proposition 3. *A completely connected network g is pairwise stable if and only if:*

1. *The linking cost is relatively low.*

$$c < \min\{\Delta^B d, \, \Delta^A d, \, \gamma_A^1, \, \gamma_B^1\}$$

   where:

$$\gamma_k^1 = \frac{\Delta^k s \Delta^k d - \tilde{\Delta}^k d \tilde{\Delta}^k s}{\Delta^k s + \Delta^k d - \tilde{\Delta}^k d - \tilde{\Delta}^k s}, \qquad \Delta^k s \Delta^k d > \tilde{\Delta}^k d \tilde{\Delta}^k s \quad \forall \, k \in \{A, B\}$$

2. *The intensity of homophily of every agent does not take extreme values.*

$$\underline{\lambda}^k \, \bar{\lambda}^k \in (a_k, \, b_k) \qquad \forall k \in \{A, B\}$$

where:

$$a_k = \frac{c - \tilde{\Delta}^k d}{\Delta^k s - \tilde{\Delta}^k d}, \qquad b_k = \frac{c - \Delta^k d}{\tilde{\Delta}^k s - \Delta^k d} \quad 0 \leq a_k < b_k \leq 1 \qquad \forall k \in \{A, B\}$$

This result is straightforward to follow, because the linking cost is low and individuals do not have an extreme preference of linking with similar or dissimilar individuals (the intensity of homophily is close to one half). When the linking cost is not so low, individuals do not link with all the agents. In fact, they will link more with similar or different individuals according to their homophilous bias in their preferences. When the intensity of homophily of some individual is beyond its upper bound, the individual has homophilous preferences and he breaks a link with a dissimilar individual. Similarly, when the intensity of homophily of some individual is less than its lower bound, then this individual breaks a link with a similar agent.

I think that completely connected networks could emerge in small indigenous communities, where individuals of different clans or groups need to communicate and cooperate among them to survive. It is plausible that these individuals do not have extreme homophilous preferences because they belong to the same community. The linking cost is relatively low, because these communities are small and they can be connected with the rest of individuals.

Finally, the bias in the networks found in Patacchini and Zenou (2016) is reproduced in this thesis (see the figure 1 in Chapter 1). For that reason, in this thesis will be used the inbreeding homophily concept. This name was given by Coleman and is also known as the Coleman index. This measure was used by Currarini et al. (2009) and it quantifies the homophilous bias that an individual presents.

For an individual $i \in N$ of type $k \in \{A, B\}$, his inbreeding homophily in the network g is computed as:

$$H_i(g) = \frac{r_s^i(g) - w_i}{1 - w_i},$$

where the relative fraction of type k in the population is represented as $w_i = \frac{n_k}{n}$. If the Coleman index is close to one, the individual shows an homophilous bias. If this measure is close to

minus one, the individual shows a more heterophilous bias. When the index is zero, this means that the individual have baseline homophily because $r_s^i(g) = w_i$.[6]

In the segregated network in dyads, all individuals have a Coleman index of one. If all individuals have the same intensity of homophily and this intensity decreases enough, then each individual makes a new crossed link. If $n_A = n_B$, the Coleman index of each individual changes from one to zero. Then, all individuals stop being so homophilous in their links and they have baseline homophily.

Similarly, all individuals have the same index in the completely connected network. This index is $H_i = -\frac{1}{n-1} \quad \forall i \in N$. If all individuals have the same intensity of homophily and this intensity increases and becomes more homophilous, then the Coleman index of each individual of type k also increases and it is $H_i = -\frac{1}{n-2} + \frac{n_k}{(n-2)(n_{-k})} \quad \forall i \in N_k, \forall k \in \{A, B\}$. When the intensity of homophily decreases, the Coleman index of each individual also decreases to $H_i = -\frac{2}{n-2} \quad \forall i \in N_k, \forall k \in \{A, B\}$.

Notice that all individuals have the same Coleman index in each of the previous networks, and they do not represent the bimodality of the inbreeding homophily in the left panel of figure 1, where some individuals of certain type show homophily in their links and other individuals show heterophily in their links.

I want to characterize a similar network to figure 3. In this network, there are two types of individuals A and B. In this figure, individuals A are homophilous in their links. However, two individuals of type B are heterophilous and four individuals are homophilous. Heterophilous individuals are represented as B^* and homophilous individuals are presented as B^{**}.

---

[6]Notice that this measure has a slight bias in small samples, because numerator of r_s^i does not take into account the individual i in the numerator whereas w_i does. See Currarini et al. (2009) & Coleman (1958) for a complete exposition of the index and its applications.
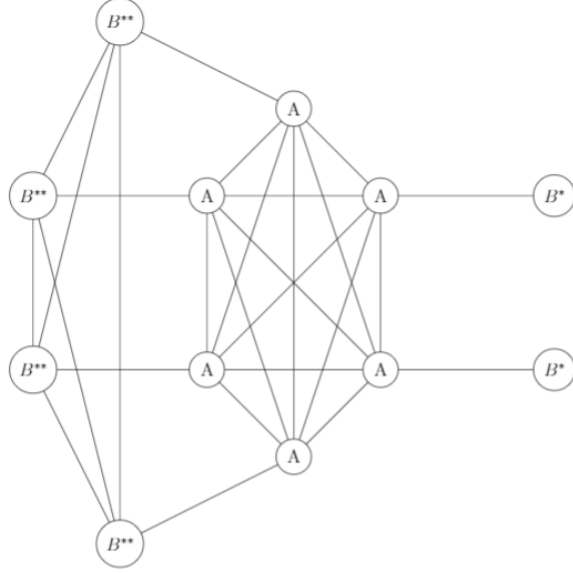
Figure 3: Network with bimodality in Coleman's index for type B. Source: Own elaboration.

The indviduals of type B are splited into two groups: the group of B^* individuals is conformed by heterophilous individuals and the group of B^{**} individuals is conformed by homophilous agents. Computing the Coleman index per group:[7]

$$H_A = \frac{\frac{5}{6} - \frac{6}{12}}{1 - \frac{6}{12}} = \frac{2}{3}$$

$$H_{B^*} = \frac{0 - \frac{6}{12}}{1 - \frac{6}{12}} = -1$$

$$H_{B^{**}} = \frac{\frac{3}{4} - \frac{6}{12}}{1 - \frac{6}{12}} = \frac{1}{2}.$$

Other scientists have found stable networks, exhibiting a bimodality in the Coleman indexes. Using the connections model, De Marti and Zenou (2017) show the two examples of stable networks that exhibit the main features of bimodality and unimodality of figure 1. The next figure shows the examples proposed by De Marti and Zenou. In the left panel, the index is unimodal for types A and bimodal for types B. In the right panel, the index is bimodal for both types.

---

[7]It is represented the Coleman index *per group* because all individuals of the same group behave similarly.
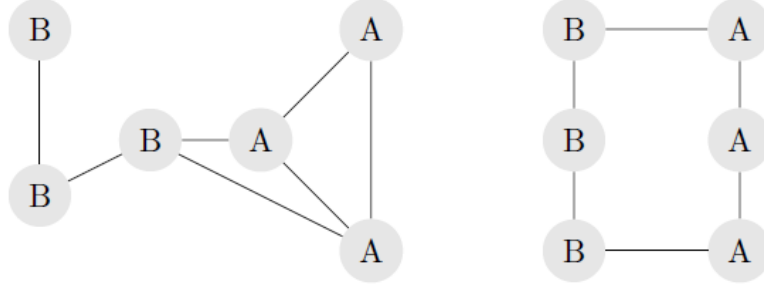
Figure 4: Stable networks proposed in De Marti and Zenou (2017) showing example 4 (left panel) and figure 9.c (right panel). Source: This figure is taken from De Marti and Zenou (2017)

It is important to remark that these examples are specific cases and there is no full characterization of pairwise stable networks similar to these examples. The next proposition brings conditions in which a network similar to figure 3 is pairwise stable.

Proposition 4. *Let* $n_A = n_B$. *A network* g *such that:*

a) *All individuals of type* A *are linked among themselves and every individual of type* A *only links with one type* B *individual.*

b) *Some individuals* B *are only linked with one type* A *individual.*

c) *There are other type* B *individuals that: (1) are linked among themselves and (2) each of them is linked to one individual of type* A,

is pairwise stable if and only if

1. *The linking cost does not take extreme values.*

$$\max\{\psi_A^2,\, 0\} < c < \min\{\gamma_A^2,\, \gamma_{B^{**}}^2,\, \Delta^A d,\, \Delta^A s,\, \Delta^{B^*} d,\, \Delta^{B^{**}} s,\, \Delta^{B^{**}} d\}$$

where:

$$\psi_A^2 = \frac{\Delta^A d\, \tilde{\Delta}^A s' - \Delta^A d'\, \tilde{\Delta}^A s}{\Delta^A d - \tilde{\Delta}^A s - \Delta^A d' + \tilde{\Delta}^A s'}$$

$$\gamma_k^2 = \frac{\Delta^k s \Delta^k d - \tilde{\Delta}^k d \tilde{\Delta}^k s}{\Delta^k s + \Delta^k d - \tilde{\Delta}^k d - \tilde{\Delta}^k s}, \quad \Delta^k s \Delta^k d > \tilde{\Delta}^k d \tilde{\Delta}^k s \qquad \forall k \in \{A,\, B^{**}\}$$

18

2. *The intensity of homophily is relatively low for those individuals of type B only connected with agents of type A. Formally, if B^* is the set of these individuals their intensity of homophily must satisfy:*

$$\bar{\lambda}^{B^*} < \min\{\frac{\Delta^{B^*}d - c}{\Delta^{B^*}d}, \ \frac{c - \tilde{\Delta}^{B^*}d'}{\Delta^{B^*}s' - \tilde{\Delta}^{B^*}d'}\}$$

3. *The intensity of homophily for the rest of individuals does not take extreme values.*

$$\max\{\frac{c - \tilde{\Delta}^A d}{\Delta^A s - \tilde{\Delta}^A d}, \frac{\Delta^A d' - c}{\Delta^A d' - \tilde{\Delta}^A s'} \ 0\} < \underline{\lambda}^A \leq \bar{\lambda}^A < \frac{\Delta^A d - c}{\Delta^A d - \tilde{\Delta}^A s}$$

$$\frac{c - \tilde{\Delta}^{B^{**}}d}{\Delta^{B^{**}}s - \tilde{\Delta}^{B^{**}}d} < \underline{\lambda}^{B^{**}} \leq \bar{\lambda}^{B^{**}} < \frac{\Delta^{B^{**}}d - c}{\Delta^{B^{**}}d - \tilde{\Delta}^{B^{**}}s}$$

Condition 1 implies that individuals can link because it is not so costly make it, but its lower bound also implies that they cannot link with so many people. This assumption seems plausible for American students, because making friends cannot be so costly but eventually people stop making new friends. Condition 2 guarantees heterophilous preferences for those individuals in the group B^*. Maybe some black students prefer to be with white students because they share similar interests or way of life. For that reason, they only link with a dissimilar individuals. Condition 3 implies that individuals have homophilous preferences, but not so homophilous because they also make a crossed link.

Notice that if individuals of the group B^* does not satisfy condition 2, they will have more homophilous preferences and it is possible that they may link with other individuals of the group B^{**}. Of course, it is also needed that individuals in the group B^{**} are willing to form these links. This could happen if thes individuals also have homophilous preferences or if the linking cost is lower. Eventually, the bimodality in Coleman's indexes of individuals of type B become a

unimodal distribution, where all individuals homophilously link.

Proposition 4 characterizes a particular network which exhibits bimodality in Coleman indexes for some type of individuals. Figure 1 represents this bimodality for black students in American colleges, whereas white students show a homophilous distribution in their indexes. Britpop musicians also show these tendencies, because some artists make professional and sentimental relationships with individuals from their same origin social class whereas there are other musicians which prefer to link with dissimilar individuals. Millward et al. (2017)

It is important to clarify that these networks are a rough representation of these phenomena, because the distribution of the Coleman indexes would be one bar for individuals of type A and two bars for individuals of type B. However, I think that this is a first approach to analyze bimodality in the Coleman indexes. In the next chapter it will be discussed some limitations of this model as well as aspects to be considered for future research.

Chapter 4

Discussion and conclusions

In this thesis, I was interested in the role of homophilous and heterophilous preferences in the emergence of stable networks. It was proposed a network formation model where individuals have homophilous preferences, they only care about the ratio of similar and dissimilar friends, and there is a constant linking cost. Finally, some networks were characterized, providing the conditions to guarantee stability.

It has been proved that a pairwise stable network cannot be segregated and fully intraconnected at the same time, because in this model individuals care about the ratio of similar and dissimilar friends. This result can be compared with the previous arguments of effective communication, because homophilous and heterophilous ties help individuals to get a new job Lin et al. (1981), Granovetter (1970) and Rogers and Bhowmik (1970). Similarly, it was proved that the only segregated network is those formed by dyads.

Finally, it was proposed a network where some individuals of the same type are homophilous and other individuals are heterophilous. The conditions to guarantee stability are intuitive, because the individuals who had heterophilous preferences showed negative Coleman indexes, whereas the rest of homophilous individuals had positive Coleman indexes. These conditions are the main result, because I have characterized a network which exhibits a bimodality in the Coleman indexes that Patacchini and Zenou (2016) reported.

The most controversial assumption is related to the homophilous and heterophilous prefer-ences. If individuals are homophilous, they will link with similar agents. I recognize that this model is a first approach, where individuals have these preferences somehow. It will be inter-esting to propose another model where an underlying process produces these homophilous and heterophilous preferences. For instance, in Bagues and Perez-Villadoniga (2013) a similar-to-me-in-skills effect arises because evaluators highly rank those individuals with similar abilities. Another example can be found in Kets and Sandroni (2019), where the uncertainty of individ-uals' actions leads to homophilous interactions because players face less strategic uncertainty when they interact with members of their same group.

For future research, it could be proposed and make some comparative statics of more pair-wise stable networks, where certain networks transit from some configuration to others as the intensity of homophily and heterophily change. A welfare analysis could also be made, finding those networks which are the most efficient in a utilitarian sense.

Appendix A

Proofs of the propositions

Proposition 1. If a fully intraconnected network is pairwise stable, then it cannot be segregated.

Proof of proposition 1. Suppose the contrary, that is there is a network g which is fully intraconnected and segregated. Because g is segregated, for all individual i in the network, $r_s^i(g) = \frac{n_i(g)}{n_i(g)} = 1$ and $r_d^i(g) = 0$. Because g is fully intraconnected, pairwise stability implies that:

$$u_i(g) = \lambda_i f(1) - c\, n_i(g) \geq \lambda_i f(1) - c\,(n_i(g) - 1) = u_i(g - ij)$$

$$\Longrightarrow$$

$$0 \geq c.$$

But this is a contradiction, because it was supposed $c > 0$.

$\square$

Before the proof of the proposition 2, I will call individuals of certain types as his type when there is no possibility of confusion. For example, if individual i is such as $i \in N_A$, then $i \equiv A$. A crossed link is represented as AB, a link between two individuals of type A is represented as AA and a link between two individuals of type B is represented as BB.

Proposition 2. Let n_B even. The only segregated pairwise stable network is that in which individuals link in dyads (as the one in panel 2b). The network in dyads is pairwise stable if and

only if the linking cost is intermediate ($\max_{i \in N_k}\{0, f(1/2) - \lambda_i f(1)\} \leq c \leq \min_{i \in N}\{\lambda_i f(1)\}$ for some $k \in \{A, B\}$) and preferences are sufficiently homophilous ($\lambda_k \geq \frac{f(1/2)}{2 f(1)}$ for the previous $k \in \{A, B\}$). Thus, no individual breaks his same type link and even the more heterophilous individuals do not want to make a crossed link.

Proof of proposition 2. Let n_B even. Supposing the contrary, that is there is a segregated pairwise stable network and it is different of the network in dyads. This means that some individual is linked with three or more similar agents. Without loss of generality, If it is supposed that the type of this individual is A and he is linked with m_A individuals, where $m_A \geq 3$.[1] But this individual is better if he removes a link, because $u_A(g - AA) \geq u_A(g)$. This is due to:

$$u_A(g - AA) = \lambda_A f(1) - c\,(m_A - 1) \geq \lambda_A f(1) - c\,m_A \qquad \text{for} \quad m_A \geq 3.$$

For that reason, the only segregated pairwise stable network is the network in dyads. Now, I have to prove to conditions that guarantee that this network is pairwise stable.

Let's suppose that g is a network in which individuals link in dyads. I have to prove that each individual does not want to break his only link and does not want to add new links. Without loss of generality, let's take an individual of type A. It means that he links with a similar individual, this implies $r_d^A(g) = \frac{0}{1} = 0$ and $r_s^A(g) = 1$. His utility function is $u_A(g) = \lambda_A f(1) - c$

- Breaking a link

I have to prove that $u_A(g) \geq u_A(g - AA)$. Suppose that the linking cost is not so high. This condition can be expressed as:

$$c \leq \min_{i \in N}\{\lambda_i f(1)\} \tag{A.1}$$

But this is exactly the condition that guarantees that the individual does not want to break his

---

[1]It will be used the notation m_A instead of n_A because it could be confused with the cardinality of the set of individuals of type A.

link, because

$$u_A(g) = \lambda_A f(1) - c \geq 0 = u_A(g - AA)$$

- Adding a link

Notice that individuals do not add a new link with another similar agent, because a new link does not add more benefits and they pay twice the linking cost. I have to prove that crossed links are not possible. The next condition guarantees this result:

$$\max\{0,\ f(1/2) - \lambda_i f(1)\} < c \qquad \forall i \in N_k \qquad \text{for some} \quad k \in \{A, B\} \qquad \text{(A.2)}$$

$$\underline{\lambda_k} \geq \frac{f(1/2)}{2\,f(1)} \qquad \text{for some} \quad k \in \{A,\ B\} \qquad \text{(A.3)}$$

This condition can be interpreted as the linking cost and their intensity of homophily are relatively high for some type of individuals. In particular, the condition A.3 guarantees that conditions A.1 and A.2 are compatible.

I prove that condition A.2 is also a necessary and sufficient condition to guarantee that crossed links will not be added. The condition A.2 is interpreted as: for at least one type of individuals, the linking cost of adding a new crossed link is higher than the benefit of this new link. If this condition does not hold, there are at least one individual in of each type such as they do not satisfy the condition A.2. Then, these individuals would make a new link between them and the network g would not be pairwise stable and segregated. $\qquad\square$

Proposition 3. A completely connected network g is pairwise stable if and only if:

1. The linking cost is relatively low.

2. The intensity of homophily of every agent does not take extreme values.

Proof of proposition 3. Let's suppose that g is a completely connected network. This implies:

$$n_i(g) = n - 1, \quad s_i(g) = n_k - 1, \quad d_i(g) = n_{-k}, \qquad \forall i \in N_k,\ \forall k \in \{A, B\}$$

25

In this proposition, the linking cost is relatively low means:

$$c < \min\{\Delta^B d,\ \Delta^A d,\ \gamma_A^1,\ \gamma_B^1\} \tag{A.4}$$

where:

$$\gamma_k^1 = \frac{\Delta^k s \Delta^k d - \tilde{\Delta}^k d \tilde{\Delta}^k s}{\Delta^k s + \Delta^k d - \tilde{\Delta}^k d - \tilde{\Delta}^k s}, \qquad \Delta^k s \Delta^k d > \tilde{\Delta}^k d \tilde{\Delta}^k s \quad \forall\, k \in \{A, B\}$$

The last condition related to the intensity of homophily is expressed as:

$$\underline{\lambda}^k\ \bar{\lambda}^k \in (a_k,\ b_k) \qquad \forall k \in \{A, B\} \tag{A.5}$$

where:

$$a_k = \frac{c - \tilde{\Delta}^k d}{\Delta^k s - \tilde{\Delta}^k d}, \qquad b_k = \frac{c - \Delta^k d}{\tilde{\Delta}^k s - \Delta^k d} \quad 0 \le a_k < b_k \le 1 \qquad \forall k \in \{A, B\}$$

Because everybody is linked in this network, I have to verify that for every individual does not want to break their links. Without loss of generality, it will be supposed that $i \in N$ is an individual of type A. I will prove that $u_A(g) \ge u_A(g - AB)$ and $u_A(g) \ge u_A(g - AA)$.

- The individual does not want to break a link with an individual of type A.

I have to prove

$$u_A(g) \ge u_A(g - AA)$$

If this inequality is satisfied, then

$$u_A(g) - u_A(g - ij) \ge 0$$

$$\lambda_A\, \Delta^A s + (1 - \lambda_A)\, \tilde{\Delta}^A d - c \ge 0$$

$$\lambda_A \left(\Delta^A s - \tilde{\Delta}^A d\right) + \tilde{\Delta}^A d - c \ge 0$$

Because $\Delta^A s - \tilde{\Delta}^A d > 0$, the last inequality can be expressed as:

$$\lambda_A \geq \frac{c - \tilde{\Delta}^A d}{\Delta^A s - \tilde{\Delta}^A d}$$

But this mathematical condition is satisfied because it has been supposed condition A.5, that is:

$$\lambda_A \geq \underline{\lambda}^A \geq \frac{c - \tilde{\Delta}^A d}{\Delta^A s - \tilde{\Delta}^A d}$$

this means that $u_A(g) \geq u_A(g - AA)$.

- The individual does not want to break a link with an individual of type B.

I have to prove

$$u_A(g) \geq u_A(g - AB)$$

Substituting these utilities

$$u_A(g) - u_A(g - AB) \geq 0$$

$$\lambda_A \, \tilde{\Delta}^A s + (1 - \lambda_A) \, \Delta^A d - c \geq 0$$

$$\lambda_A \left(\tilde{\Delta}^A s - \Delta^A d \right) + \Delta^A d - c \geq 0$$

because $\tilde{\Delta}^A s - \Delta^A d < 0$, this last inequality can be written as:

$$\lambda_A \leq \frac{c - \Delta^A d}{\tilde{\Delta}^A s - \Delta^A d}$$

But it was supposed the condition A.5 and this implies:

$$\lambda_A \leq \bar{\lambda}^A \leq \frac{c - \Delta^A d}{\tilde{\Delta}^A s - \Delta^A d}$$

Because this inequality holds, the individual does not have incentives to break his link with

individual B.

$$u_A(g) \geq u_A(g - AB)$$

I can prove that every individual of type B does not want to break his links. For that reason, I have proved that a completely connected network g is pairwise stable.

$\square$

Notice that condition A.4 is made to guarantee $0 \leq a_k < b_k \leq 1 \quad \forall k \in \{A, B\}$. Condition A.4 implies that the marginal benefit of linking with all individuals must be greater than its marginal cost.

Proposition 4. Let $n_A = n_B$. A network g such that:

i) Every individual $i \in N_A$ is connected with the rest of individuals in N_A and he is linked with one individual B.

ii) Some individuals $i \in N_B$ are only linked with one individual A. The set of these individuals is called N_{B*} and its cardinality n_{B*}

iii) Some individuals $i \in N_B$ are linked with one individual A and the rest of individuals in N_B, excepting those individuals that satisfy $ii)$. The set of this individuals is called N_{B**} and its cardinality n_{B**},

is pairwise stable if and only if

1. The linking cost does not take extreme values.

2. The intensity of homophily is relatively low for those individuals of type B only connected with agents of type A.

3. The intensity of homophily for the rest of individuals does not take extreme values.

Proof of proposition 4. Suppose g is a network that satisfies conditions $i)$, $ii)$ and $iii)$ of Proposition 4. Because all individuals have one crossed link, this means that $d_i(g) = 1 \quad \forall i \in N$ and:

$$s_i(g) = \begin{cases} n_A - 1 & \forall i \in N_A \\ 0 & \forall i \in N_{B^*} \\ n_{B^{**}} - 1 & \forall i \in N_{B^{**}} \end{cases} \qquad n_i(g) = \begin{cases} n_A & \forall i \in N_A \\ 1 & \forall i \in N_{B^*} \\ n_{B^{**}} & \forall i \in N_{B^{**}} \end{cases}$$

I have to prove that for every $i \in N$, they do not want to remove their links and they do not prefer to add a new link that does not exist in g.

Breaking links

- Individuals of type A

Consider any individual of type A. Because he links with all individuals of type A and he links with one individual of type B, I have to verify that he prefers these links.

I have to prove that individual A does not want to break his links with another individual of type A, i.e

$$u_A(g) \geq u_A(g - AA)$$

Computing these utilities:

$$\lambda_A f\left(\frac{n_A - 1}{n_A}\right) + (1 - \lambda_A) f\left(\frac{1}{n_A}\right) - c\,n_A \geq \lambda_A f\left(\frac{n_A - 2}{n_A - 1}\right) + (1 - \lambda_A) f\left(\frac{1}{n_A - 1}\right) - c\,(n_A - 1)$$

Using the definitions of the benefit of linking with a similar individual and its associated cost:

$$\lambda_A \Delta^A s + (1 - \lambda_A)\,\tilde{\Delta}^A d \geq c$$

$$\lambda_A \left(\Delta^A s - \tilde{\Delta}^A d\right) \geq c - \tilde{\Delta}^A d$$

because the expresion $\Delta^A s - \tilde{\Delta}^A d > 0$, the last inequality can be expressed as:

$$\lambda_A \geq \frac{c - \tilde{\Delta}^A d}{\Delta^A s - \tilde{\Delta}^A d}$$

The next condition guarantees that all individuals A does not want to break a link with similar individuals.

$$\underline{\lambda}_A \geq \frac{c - \tilde{\Delta}^A d}{\Delta^A s - \tilde{\Delta}^A d} \tag{A.6}$$

Now, I have to prove that individual A does not want to break his only link with individual of type B, i.e.

$$u_A(g) \geq u_A(g - AB)$$

Computing these utilities:

$$\lambda_A f\left(\frac{n_A - 1}{n_A}\right) + (1 - \lambda_A)f\left(\frac{1}{n_A}\right) - c\,n_A \geq \lambda_A f\,(1) - c\,(n_A - 1)$$

$$\lambda_A \tilde{\Delta}^A s + (1 - \lambda_A)\Delta^A d \geq c$$

$$\lambda_A\left(\tilde{\Delta}^A s - \Delta^A d\right) \geq c - \Delta^A d$$

$$\Delta^A d - c \geq \lambda_A\left(\Delta^A d - \tilde{\Delta}^A s\right)$$

The individual of type A does not want to break his crossed link if and only if:

$$\frac{\Delta^A d - c}{\Delta^A d - \tilde{\Delta}^A s} \geq \lambda_A$$

The next condition guarantees that all individuals of type A do not want to break their crossed links.

$$\frac{\Delta^A d - c}{\Delta^A d - \tilde{\Delta}^A s} \geq \bar{\lambda}_A \tag{A.7}$$

- Individuals of type B

Because individuals of type B split into two groups, consider these possible cases.

Case 1. *The individual of type B belongs to group N_{B^*}. I have to prove that he does not*

30

want to break his only link, i.e.

$$u_{B^*}(g) \geq u_{B^*}(g - AB^*) = 0$$

$$(1 - \lambda_{B^*})f(1) - c \geq 0$$

Because $\Delta^{B^*}d = f(1)$, then

$$\frac{\Delta^{B^*}d - c}{\Delta^{B^*}d} \geq \lambda_{B^*}.$$

The next condition guarantees that individuals of the group N_{B^*} do not want to break their crossed links.

$$\frac{\Delta^{B^*}d - c}{\Delta^{B^*}d} \geq \bar{\lambda}_{B^*}. \tag{A.8}$$

Case 2. *The individual of type B belongs to group $N_{B^{**}}$.* I have to prove that an individual of the group B^{**} does not want to break his links with an individual of type A

$$u_{B^{**}}(g) \geq u_{B^{**}}(g - AB^{**})$$

$$\lambda_{B^{**}}f\left(\frac{n_{B^{**}} - 1}{n_{B^{**}}}\right) + (1 - \lambda_{B^{**}})f\left(\frac{1}{n_{B^{**}}}\right) - c\,n_{B^{**}} > \lambda_{B^{**}}f(1) - c(n_{B^{**}} - 1),$$

this condition is satisfied if and only if

$$\frac{\Delta^{B^{**}}d - c}{\Delta^{B^{**}}d - \tilde{\Delta}^{B^{**}}s} > \lambda_{B^{**}}.$$

The next condition guarantees that all individuals of the group $N_{B^{**}}$ do not want to break their crossed links.

$$\frac{\Delta^{B^{**}}d - c}{\Delta^{B^{**}}d - \tilde{\Delta}^{B^{**}}s} > \bar{\lambda}_{B^{**}}. \tag{A.9}$$

Now, I will show that individuals does not want to break their links with other members of

31

the group B^{**}. I have to prove:

$$u_{B^{**}}(g) \geq u_{B^{**}}(g - B^{**}B^{**})$$

$$\lambda_{B^{**}} f\left(\frac{n_{B^{**}} - 1}{n_{B^{**}}}\right) + (1 - \lambda_{B^{**}}) f\left(\frac{1}{n_{B^{**}}}\right) - c\,n_{B^{**}} \geq$$
$$\geq \lambda_{B^{**}} f\left(\frac{n_{B^{**}} - 2}{n_{B^{**}} - 1}\right) + (1 - \lambda_{B^{**}}) f\left(\frac{1}{n_{B^{**}} - 1}\right) - c\,(n_{B^{**}} - 1),$$

this can be written as

$$\lambda_{B^{**}} \Delta^{B^{**}} s + (1 - \lambda_{B^{**}}) \tilde{\Delta}^{B^{**}} d \geq c.$$

The B^{**} prefers to be connected with an individual of the group $N_{B^{**}}$ if and only if

$$\lambda_{B^{**}} \geq \frac{c - \tilde{\Delta}^{B^{**}} d}{\Delta^{B^{**}} s - \tilde{\Delta}^{B^{**}} d}.$$

The next condition guarantees that all individuals of the group $N_{B^{**}}$ do no want to break their links with members of the same group.

$$\underline{\lambda}_{B^{**}} > \frac{c - \tilde{\Delta}^{B^{**}} d}{\Delta^{B^{**}} s - \tilde{\Delta}^{B^{**}} d}. \tag{A.10}$$

Adding links

- Individuals of type A

Let's take an arbitrary individual of type A. Because he is linked with all the similar individual, I just have to show that he does not want to add new crossed links, i.e.

$$u_{B^{**}}(g) > u_{B^{**}}(g + AB)$$

$$\lambda_A f\left(\frac{n_A-1}{n_A}\right)+(1-\lambda_A)f\left(\frac{1}{n_A}\right)-c\,n_A > \lambda_A f\left(\frac{n_A-1}{n_A+1}\right)+(1-\lambda_A)f\left(\frac{2}{n_A+1}\right)-c\,(n_A+1).$$

Using the definition of benefits of adding a new link with a dissimilar individual:

$$c > \lambda_A \tilde{\Delta}^A s' + (1-\lambda_A)\,\Delta^A d'.$$

The condition for no addition of new crossed links is:

$$\lambda_A > \frac{\Delta^A d' - c}{\Delta^A d' - \tilde{\Delta}^A s'}.$$

If it is supposed the next condition, all the individuals of type A do not to make a new crossed link.

$$\underline{\lambda}_A > \frac{\Delta^A d' - c}{\Delta^A d' - \tilde{\Delta}^A s'}. \tag{A.11}$$

- Individuals of type B

Again, there are two possible cases for this type of individuals.

Case 1. *The individual belongs to group B^**. Because he does not obtain more utility linking with dissimilar individuals, I have to verify that he does not want to make new links with similar individuals.

$$u_{B^*}(g) > u_{B^*}(g + B^* B^{**}),$$

this expression can be written as

$$(1-\lambda_{B^*})f(1) - c > \lambda_{B^*} f(1/2) + (1-\lambda_{B^*})f(1/2) - 2\,c$$

$$c > \lambda_{B^*}\Delta^{B^*} s' + (1-\lambda_{B^*})\tilde{\Delta}^{B^*} d',$$

where $\Delta^{B^*} s' = f(1/2)$. This last inequality can be represented as:

$$\frac{c - \tilde{\Delta}^{B^*} d'}{\Delta^{B^*} s' - \tilde{\Delta}^{B^*} d'} > \lambda_{B^*}.$$

The next condition guarantees that individuals of the group N_{B^*} do not want to make links with individuals of the group $N_{B^{**}}$.

$$\frac{c - \tilde{\Delta}^{B^*} d'}{\Delta^{B^*} s' - \tilde{\Delta}^{B^*} d'} > \bar{\lambda}_{B^*}. \tag{A.12}$$

Case 2. *The individual belongs to group B^{**}.* Notice that an individual of the group B^{**} cannot make new links with other individuals, because he is completely linked with the members of his same group, whereas individuals of the group B^* and type A individuals do not want to make new links.

As it was shown, the network g is pairwise stable because all individuals of different types are better with their links than without deleting each of them at a time, and they do not form new links.

The following conditions A.13 to A.16 are assumed to make compatible the conditions A.6 to A.12. If these conditions hold, then the network g is pairwise stable.

$$\max\{\psi_A^2, \, 0\} < c < \min\{\gamma_A^2, \, \gamma_{B^{**}}^2, \, \Delta^A d, \, \Delta^A s, \, \Delta^{B^*} d, \, \Delta^{B^{**}} s, \, \Delta^{B^{**}} d\} \tag{A.13}$$

where

$$\psi_A^2 = \frac{\Delta^A d \, \tilde{\Delta}^A s' - \Delta^A d' \, \tilde{\Delta}^A s}{\Delta^A d - \tilde{\Delta}^A s - \Delta^A d' + \tilde{\Delta}^A s'}$$

$$\gamma_k^2 = \frac{\Delta^k s \Delta^k d - \tilde{\Delta}^k d \tilde{\Delta}^k s}{\Delta^k s + \Delta^k d - \tilde{\Delta}^k d - \tilde{\Delta}^k s}, \quad \Delta^k s \Delta^k d > \tilde{\Delta}^k d \tilde{\Delta}^k s \qquad \forall k \in \{A, \, B^{**}\}$$

The intensity of homophily for those individuals of type B who belong to group N_{B^*} satisfies:

$$\bar{\lambda}^{B^*} < \min\{\frac{\Delta^{B^*} d - c}{\Delta^{B^*} d}, \, \frac{c - \tilde{\Delta}^{B^*} d'}{\Delta^{B^*} s' - \tilde{\Delta}^{B^*} d'}\} \tag{A.14}$$

The intensity of homophily for the rest of individuals satisfies the next two conditions:

$$\max\{\frac{c - \tilde{\Delta}^A d}{\Delta^A s - \tilde{\Delta}^A d}, \frac{\Delta^A d' - c}{\Delta^A d' - \tilde{\Delta}^A s'} \; 0\} < \underline{\lambda}^A \leq \bar{\lambda}^A < \frac{\Delta^A d - c}{\Delta^A d - \tilde{\Delta}^A s} \qquad (A.15)$$

$$\frac{c - \tilde{\Delta}^{B^{**}} d}{\Delta^{B^{**}} s - \tilde{\Delta}^{B^{**}} d} < \underline{\lambda}^{B^{**}} \leq \bar{\lambda}^{B^{**}} < \frac{\Delta^{B^{**}} d - c}{\Delta^{B^{**}} d - \tilde{\Delta}^{B^{**}} s} \qquad (A.16)$$

Notice that condition A.13 assumes a similar condition to the condition A.4 in proposition 2, where the marginal benefits of linking for individuals of type A and B^{**} are greater than their associated costs.

The condition A.13 mens that the linking cost does not take extreme values. The condition A.14 means that the intensity of homophily is relatively low for individuals of the group N_{B^*} and the conditions A.15 and A.16 means that the intensity of homophily for the rest of individuals do not take extreme values. $\qquad \square$

References

Bagues, M. and Perez-Villadoniga, M. J. (2013). Why do i like people like me? *Journal of Economic Theory*, 148(3):1292–1299.

Barranco, O., Lozares, C., and Muntanyola-Saura, D. (2019). Heterophily in social groups formation: a social network analysis. *Quality & Quantity*, 53(2):599–619.

Bramoullé, Y. and Kranton, R. (2007). Risk sharing across communities. *American Economic Review*, 97(2):70–74.

Coleman, J. S. (1958). Relational analysis: the study of social organizations with survey methods. *Human organization*, 17(4):28–36.

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

Dasgupta, P. and Goyal, S. (2009). Narrow identities. Technical report, Discussion Paper, Faculty of Economics, University of Cambridge.

De Marti, J. and Zenou, Y. (2017). Segregation in friendship networks. *The Scandinavian Journal of Economics*, 119(3):656–708.

Fafchamps, M. and Gubert, F. (2007). Risk sharing and network formation. *American Economic Review*, 97(2):75–79.

Granovetter, M. S. (1970). *Changing jobs: channels of mobility information in a suburban population*. PhD thesis, Harvard University.

Granovetter, M. S. (1977). The strength of weak ties. In *Social networks*, pages 347–367. Elsevier.

Ibarra, H. (1992). Homophily and differential returns: Sex differences in network structure and access in an advertising firm. *Administrative science quarterly*, pages 422–447.

Iijima, R. and Kamada, Y. (2017). Social distance and network structures. *Theoretical Economics*, 12(2):655–689.

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

Jha, Y. and Welch, E. W. (2010). Relational mechanisms governing multifaceted collaborative behavior of academic scientists in six fields of science and engineering. *Research Policy*, 39(9):1174–1184.

Jones, B. F., Wuchty, S., and Uzzi, B. (2008). Multi-university research teams: Shifting impact, geography, and stratification in science. *science*, 322(5905):1259–1262.

Kalmijn, M. (1998). Intermarriage and homogamy: Causes, patterns, trends. *Annual review of sociology*, 24(1):395–421.

Kerckhoff, A. C. and Back, K. W. (1965). Sociometric patterns in hysterical contagion. *Sociometry*, pages 2–15.

Kets, W. and Sandroni, A. (2019). A belief-based theory of homophily. *Games and Economic Behavior*, 115:410–435.

Korte, C. and Milgram, S. (1970). Acquaintance networks between racial groups: Application of the small world method. *Journal of Personality and Social Psychology*, 15(2):101.

Lazarsfeld, P. F., Merton, R. K., et al. (1954). Friendship as a social process: A substantive and methodological analysis. *Freedom and control in modern society*, 18(1):18–66.

Lin, N., Ensel, W. M., and Vaughn, J. C. (1981). Social resources and strength of ties: Structural factors in occupational status attainment. *American sociological review*, pages 393–405.

Little, A. C., Burt, D. M., and Perrett, D. I. (2006). Assortative mating for perceived facial personality traits. *Personality and Individual Differences*, 40(5):973–984.

Liu, W. T. and Duff, R. W. (1972). The strength in weak ties. *Public Opinion Quarterly*, 36(3):361–366.

Lozares, C., Verd, J. M., Cruz, I., and Barranco, O. (2014). Homophily and heterophily in personal networks. from mutual acquaintance to relationship intensity. *Quality & Quantity*, 48(5):2657–2670.

Marsden, P. V. (1988). Homogeneity in confiding relations. *Social networks*, 10(1):57–76.

McPherson, M., Smith-Lovin, L., and Cook, J. M. (2001). Birds of a feather: Homophily in social networks. *Annual review of sociology*, 27(1):415–444.

Millward, P., Widdop, P., and Halpin, M. (2017). A 'different class'? homophily and heterophily in the social class networks of britpop. *Cultural Sociology*, 11(3):318–336.

Moody, J. (2004). The structure of a social science collaboration network: Disciplinary cohesion from 1963 to 1999. *American sociological review*, 69(2):213–238.

Newman, M. E. (2000). Who is the best connected scientist? a study of scientific coauthorship networks. *arXiv preprint cond-mat/0011144*.

Patacchini, E. and Zenou, Y. (2016). Racial identity and education in social networks. *Social Networks*, 44:85–94.

Podolny, J. M. (1993). A status-based model of market competition. *American journal of sociology*, 98(4):829–872.

Rogers, E. M. and Bhowmik, D. K. (1970). Homophily-heterophily: Relational concepts for communication research. *Public opinion quarterly*, 34(4):523–538.

Schelling, T. C. (1971). Dynamic models of segregation. *Journal of mathematical sociology*, 1(2):143–186.

Streukens, S. and Andreassen, T. W. (2013). Customer preferences for frontline employee traits: Homophily and heterophily effects. *Psychology & Marketing*, 30(12):1043–1052.

Thiessen, D. and Gregg, B. (1980). Human assortative mating and genetic equilibrium: An evolutionary perspective. *Ethology and Sociobiology*, 1(2):111–140.

Thiessen, D., Young, R. K., and Delgado, M. (1997). Social pressures for assortative mating. *Personality and Individual Differences*, 22(2):157–164.