Segregation in Friendship Networks*

Joan de Martí[†] Yves Zenou[‡]

September 2, 2013 (Revised November 30, 2015)

Abstract

We analyze a network formation model where agents belong to different communities. Both individual benefits and costs depend on direct as well as indirect connections. Benefits of an indirect connection decrease with distance in the network while the cost of a link depends on the type of agents involved. Two individuals from the same community always face a low linking cost while the cost of forming a relationship between two individuals from different communities diminishes with the rate of exposure of each of them to the other community. We find that socialization among the same type of agents can be weak even if the cost of maintaining links within one's own type is very low. Our model also suggests that policies aiming at reducing segregation are socially desirable only if they reduce the within-community cost differential by a sufficiently large amount.

Keywords: social networks, segregation, homophily, social norms.
JEL Classification: J15, Z13.

*We are grateful to the editor Matti Liski and two anonymous referees for very helpful comments. We also thank the participants of the 13th Coalition Theory Network Workshop, of the seminars at IFN, IIES, and Duke University, in particular, Andrea Galeotti, Guillaume Haeringer, Rachel Kranton, and Karl Schlag for helpful comments, and to Willemien Kets for a thorough discussion. Joan de Marti gratefully acknowledges financial support from the Spanish Ministry of Education through the project SEJ2006-09993/ECON and a Juan de la Cierva fellowship, co-financed by the European Social Fund, and the support of Barcelona GSE and the Government of Catalonia. Yves Zenou gratefully acknowledges the financial support from the French National Research Agency grant ANR-13-JSH1-0009-01.

[†]Universitat Pompeu Fabra and Barcelona GSE. E-mail: joan.demarti@upf.edu.

[‡]Stockholm University, Research Institute of Industrial Economics (IFN), Sweden, and CEPR. E-mail: yves.zenou@ne.su.se.

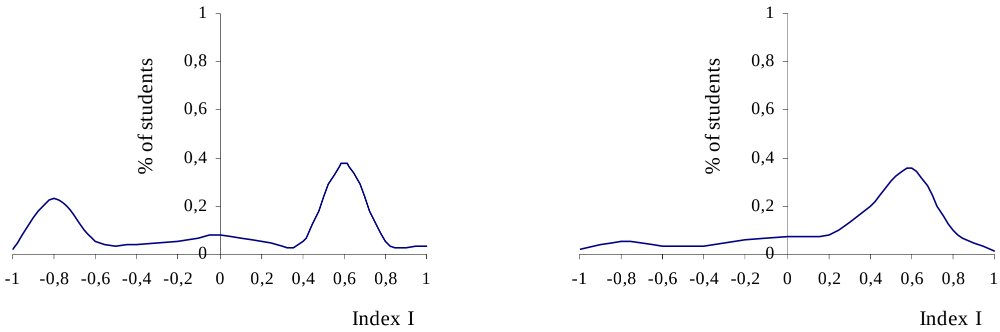
1 Introduction

In social and economic contexts, agents generally come with relevant attributes, such as ethnicity, gender, age, education, income, etc., and those attributes are often related to the interaction pattern. Are individuals more likely to be linked to others who have similar characteristics? This is a phenomenon known as *homophily*, and it refers to the fairly pervasive observation in working with social networks that having similar characteristics is often a strong and significant predictor of two individuals being connected (McPherson et al., 2001). This means that social networks can, and often do, exhibit strong segregation patterns. Segregation can occur because of the decisions of the people involved and/or by forces that affect the ways in which they meet and have opportunities to interact (Currarini et al., 2009, 2010; Tarbush and Teytelboym, 2014). Clearly, capturing homophily requires one to model or at least explicitly account for characteristics of nodes that exhibit a dimension of heterogeneity across the population.

The aim of this paper is to develop a network-formation model where agents are heterogeneous in some observable characteristics (such as ethnicity), which imply different interaction costs between communities, and where homophily behavior and segregation emerge in equilibrium.

Consider, for example, Figure 1, taken from Patacchini and Zenou (2016), which depicts a friendship network among high school students in the United States (from the National Longitudinal Survey of Adolescent Health – 'AddHealth'). It turns out that the (self-reported) friendships are strongly related to ethnicity, with students of the same ethnicity being significantly more likely to be connected to each other than students of different ethnicities. To be more precise, Patacchini and Zenou (2016) use the homophily index H_i of individual i proposed by Coleman (1958) to analyze the exposure of individuals of white and black race to own and other races. If the homophily index H_i of a student i is equal to 0, it means that the percentage of same-race friends of this individual equals the share of same-race students in the school. Negative values of the index imply an underexposure to same race students, while positive values imply an overexposure to same race students compared to the mean. Figure 1 displays their results for mixed schools (i.e. schools with a percentage of black and white students between 35 and 75 percent). Most of white students have white friends since roughly 40 percent of them are associated with values of the homophily index greater than 0.4, denoting a clear deviation from the assumption of random choice of friends by race. Black students appear to be more heterogenous in their choice of friends than whites. The clear bimodality in the distribution (corresponding to values of H_i between -0.6 and -0.8 and between 0.6 and 0.8) reveals that there are, mainly, two types of black students: those who have mostly white friends and those choosing mostly black friends.[1]

[1] Marmaros and Sacerdote (2006) and Patacchini et al. (2015) show that the main determinants of friendship formation are the geographical proximity and race. Also Mayer and Puller (2008), using administrative data and information from Facebook.com, find that race is strongly related to social ties, even after controlling for a variety of measures of socioeconomic background, ability, and college activities.



Blacks in integrated schools Whites in integrated schools

Figure 1. Distribution of students by share of same-race friends in integrated schools

In this paper, we propose a network formation model that can explain the socialization patterns observed in Figure 1. For that, we consider a finite population of individuals composed of two different communities. These two communities are categorized according to some exogenous factor such as, for example, their gender, race or ethnic and cultural traits. Individuals decide with whom they want to form a link according to a utility function that weights the costs and benefits of each connection. This results in a network of relationships where a link between two different individuals represents a friendship relationship. The utility of each individual depends on the geometry of this friendship network.

To model the benefits and costs of a given network, we consider a variation of the connections model introduced by Jackson and Wolinsky (1996), a workhorse model in the analysis of strategic network formation.[2] From the standard connections model, we keep the property that an individual benefits from her direct and indirect connections, and that this benefit decays with distance in the network. This can be interpreted as positive externalities derived from information transmission (of trends and fashion for adolescents, of job offers for workers, etc.). However, in the standard connections model, each link is equally costly, irrespective of the pair of agents that is connected. We depart from this assumption as follows.

Consider the case where communities are defined according to ethnicity, which may entail differences in language and social norms. When two individuals of different communities interact, they may initially experience a disutility due to the attachment to their original culture. This discomfort can, however, be mitigated if individuals are frequently exposed to the other community. Indeed, when someone spends time interacting with people from the other community, she can learn the codes and norms (prescriptions) that govern their social interactions. This is precisely the starting point of our analysis: the *exposure* to another social group decreases the cost of interacting with individuals from that group.

[2] See Goyal (2007), Jackson (2008), de Martí and Zenou (2011) and Jackson and Zenou (2015) for overviews of the growing literature on social and economic networks.

To be more precise, we assume that the linking cost of a pair of agents belonging to different communities depends on their level of exposure to the other community. We model this feature through a cost function that *positively* depends on the fraction of same-type friends each person has. This cost is, however, never lower than the intra-community linking cost.

In this respect, social distance expresses the force underlying this cost structure. Two agents are closer in the social space, the more each of them is exposed to the other community. And, the closer they are in the social space, the easier it is for them to interact. In our model, *this social distance is endogenous* and depends on the respective choice of peers.

We study the shape of stable networks in this setup. We use the notion of pairwise stability, introduced by Jackson and Wolinsky (1996). It is a widespread tool in the strategic analysis of social and economic networks. It takes into account the *individual incentives* to create and sever links and the necessary *mutual consent* between both sides for a link to be formed. In a nutshell, a network is pairwise stable if no agent has incentives to sever any of her links, and no pair of agents who are not connected have incentives to form a new link. In our model, it is a complex combinatorial problem to fully characterize the set of stable networks. We provide, however, a partial characterization that conveys information about the different socialization patterns that may arise in equilibrium.

In this context, when *intra-community* linking costs are low, we show that two communities may be integrated or segregated depending on the *inter-community* costs. We also show, that, in several equilibrium configurations, *bridge links* (i.e. links that connect both communities) prevail. Even if those bridge links can be quite costly for the agents involved, these links give them direct access to parts of the networks that would be not accessible otherwise. This reverberates into direct and indirect benefits that overcome the cost for both sides of the link, and acts as a positive externality for the agents who are in their respective neighborhoods since the cost of a link is only paid by the individuals directly involved.

The mechanism we suggest links *socialization costs* with *network geometry*. Since individual and aggregate welfare depends on the geometry of the resulting network, we may wonder about the impact of policies aiming at reducing inter-community socialization costs. In our context, such an analysis is difficult to perform due to the inherent multiplicity of stable configurations. We try, however, to perform one step in this direction by comparing two extreme outcomes: extremely integrated and segregated networks. When intra-community costs are low, we show that social integration is not always preferred to social segregation. The inefficiency comes from the excessive individual cost paid in building bridge links between communities. This suggests that these types of policies may only be effective if they substantially reduce inter-community socialization costs. We believe that this is an interesting result that may explain part of the relative inefficiency of integration policies such as school busing, forced integration of public housing, and Moving to Opportunity (MTO), implemented in the United States (the latter relocates families from high-to low-poverty neighborhoods (and from racially segregated to mixed neighborhoods).[3] In our theoretical framework, policies reducing inter-community socialization costs are not necessarily going to induce more desirable network structures. For example, activities outside the classroom for adolescents or cultural activities at the neighborhood level can favor integrated patterns since they may facilitate interactions among individuals of different types, but the outcome is not going to be socially efficient unless these policies sufficiently decrease the cost of interactions.

Related Literature. The papers by Currarini et al. (2009, 2010), Bramoullé et al. (2012), and Mele (2010) study homophily in networks using models of network formation. The aim in these papers is therefore similar, but there are important differences with respect to the methodology. They assume a dynamic and stochastic matching sequence while we study strategic linking decisions in a one-shot game. The papers by Currarini et al. (2009, 2010) develop a matching model with a population formed by communities of different sizes. They are able to replicate a number of observations from real-world data related to homophilous behavior at the aggregate level but, in their model, individuals' behavior is totally homogeneous within the same group of agents. Bramoullé et al. (2012) depart from Currarini et al. (2009, 2010) by assuming that dynamic matching follows the process studied by Jackson and Rogers (2007) and they show that more connected individuals tend to have a more diverse set of friends. Mele (2010) studies a model where meetings are dynamic and stochastic and each individual involved in a meeting can decide whether she wants to create or sever the link with the other person. Mele shows that this process always converges to a unique steady-state distribution.

The papers by Johnson and Gilles (2000) and Jackson and Rogers (2005) extend the Jackson and Wolinsky (1996) connection model by introducing ex ante heterogeneity in the cost structure. In the latter model, the cost of creating links between the two communities is exogenous and does not depend on the behavior of the two agents involved in the connection. In the former model, the cost of creating a link is proportional to the geographical distance between two individuals and thus this cost is fixed ex ante and does not change with the linking decisions of the two agents involved in the link. This turns out to be a key difference with our cost structure, where the cost of a link is endogenous and depends on the neighborhood structure of the two agents involved in the link. Gallo (2012) also proposes a network-formation model with ex ante heterogeneity between individuals, where the heterogeneity stems from the fact that agents have different knowledge of the network. He shows that equilibrium networks display small-world properties with segregation patterns.

Some papers analyze the consequences of homophily in social networks. For example, Golub and Jackson (2012) study how homophilous networks affect communication and agents' beliefs in a dynamic information transmission process.

Finally, Schelling (1971) is a seminal reference when discussing social networks and segregation

[3] See Lang (2007), which gives a very nice overview of these policies in the United States.

patterns. Shelling's model shows that, even a mild preference for interacting with people from the same community can lead to large differences in terms of location decision. Indeed, his results suggest that total segregation persists even if most of the population is tolerant about heterogeneous neighborhood composition.[4] Our analysis differs from Schelling's classical framework (and its different extensions) in several ways. First of all, we analyze a network formation game while in Schelling the grid of the system is fixed, but agents can choose to change their own position in the network. Secondly, homophilous preferences in our setup are not homogenous and are endogenous. In particular, these preferences are determined by both the direct and indirect benefits derived from the creation of a link and by the social environment of the potential partner. The economic benefits thus depend on the network structure of all the population.

Our main contribution. Our main contribution is to show that the mechanism of our model (that relates the cost of friendship to the social distance of two linked individuals) can induce endogenous asymmetric socialization behaviors of a particular, and economically relevant, type. We assume that socialization costs depend on exposure to other communities and we show that *ex ante identical individuals* may end up with very different network positions. Thus, we obtain intra-group asymmetric behaviors in connectivity in a number of equilibrium networks, which allow us to rationalize the friendship patterns observed in Figure 1. We do not mean here that the result of socialization is always going to lead to segregation, but we are able to show that these patterns can emerge in some circumstances as the result of a *decentralized* process of socialization. There are also other possible equilibria where this would not occur and our direct aim *is not to provide a full characterization of the set of equilibrium networks*. Indeed, the pool of high-schools from the AddHealth data set shows a variety of real-world configurations. Therefore, it is natural that any model that wants to give reasonable microfoundations for these configurations exhibits multiplicity of equilibria. We endogenously model the structure of the network of friendship relations where not only friends, but friends of friends, and friends of friends of friends, etc. matter. Because of this feature, a problem of a combinatorial nature, also present in the classical model of Jackson and Wolinsky (1996), emerges.[5] This is why it is extremely hard, if not impossible, to provide a full-fledged characterization of all possible stable networks.[6]

[4] This framework has been modified and extended in different directions, exploring, in particular, the stability and robustness of this extreme outcome (see, for example, Mobius, 2007, Zhang, 2004, or Goffette-Nagot, et al., 2012).

[5] It is indeed well-known that non-cooperative games of network formation with nominal lists of intended links are plagued by coordination problems (Myerson, 1991; Jackson, 2008; Cabrales et al., 2011). Cooperative-like stability concepts solve them partially, but heavy combinatorial costs still jeopardize a full characterization.

[6] The existence of a plethora of equilibria in our framework is not the result of the use of a weak stability concept (in our case, pairwise stability). The use of an stronger equilibrium concept in network formation games, such as Pairwise Nash equilibria, does not seem to significantly reduce the number of equilibria: in a slightly perturbed version of the present model, we are able to show that the set of pairwise stable equilibria and the set of pairwise Nash equilibria coincide. This is available upon request.

2 The model

2.1 Individuals, communities, and networks

There is a finite population of individuals denoted by $\mathbb{N} = \{1, \ldots, n\}$. This population is divided into two communities, the *Blue* (B) and the *Green* (G) communities. Each agent belongs exclusively to one of the two communities, B or G. This initial endowment of each individual can be interpreted, for example, as the ethnicity inherited from her family. The type of individual i is denoted by $\tau(i) \in \{B, G\}$. Let n^B denote the number of B individuals in the population. Similarly, let n^G denote the number of G individuals in the population. We have that $n = n^B + n^G$. We assume, without loss of generality, that $n^B \leq n^G$.

Individuals will be connected through a social network structure. A network is represented by a graph, where each node represents an individual and a connection among nodes represents a friendship relationship between the two individuals involved. We denote a network by g, and $g_{ij} = 1$ if i is friend with j and $g_{ij} = 0$ otherwise. In our framework, friendship relationships are taken to be reciprocal, i.e. $g_{ij} = g_{ji}$ so that graphs/networks are *undirected*. We denote the link of two connected individuals, i and j, by ij. The set of i's direct contacts is: $\mathbb{N}_i(g) = \{j \neq i \mid g_{ij} = 1\}$, which is of size $n_i(g)$. The direct contacts of individual i of the same type is $\mathbb{N}_i^{\tau(i)}(g) = \{j \neq i, \tau(i) = \tau(j) \mid g_{ij} = 1\}$, and we denote the cardinality of this set by $n_i^{\tau(i)}(g)$.

In Figure 2, we present some examples of network configurations. A necessary condition for a *circle* network (left panel) is that each agent has two direct contacts. The *star-shaped network* (middle panel) has one central agent who is in direct contact with all the other peripheral agents who, in turn, are only connected to this central agent. The *complete network* (right panel) is such that each agent is in direct relationship with every other agents so that each individual i has $n - 1$ direct contacts.

A network is depicted as a set of colored nodes (Figure 3), which allows us to distinguish among members of different groups, and links that connect some or all of them. Naturally, green nodes indicate type G individuals while blue nodes refer to type B individuals. Also, a *green node* is always represented by a *circle* while a *blue node* is represented by a *square*.

We still need to introduce some more concepts associated to the connectivity of the network.[7]

There is a *path* in network g from individual i to individual j if there exists an ordered set of individuals, with i being the first one and j being the last one, such that each agent is connected to the following one according to this order.[8] Graphically, there is a path from individual i to individual j whenever an agent can travel from i to j through the links of the network. The length

[7] We use the same notations as in Jackson and Zenou (2015).

[8] Formally, a *path* p_{ij}^k of length k from i to j in the network g is a sequence $\langle i_0, i_1, ..., i_k \rangle$ of players such that $i_0 = i$, $i_k = j$, $i_p \neq i_{p+1}$, and $g_{i_p i_{p+1}} = 1$, for all $0 \leq p \leq k - 1$, that is, players i_p and i_{p+1} are directly linked in g. If such a path exists, then individuals i and j are path-connected.

of a path is the number of links involved in it. The *shortest path* between i to j is the path that involves the lowest number of links. We define the *geodesic distance* (or simply distance) between individuals i and j as the length of the shortest path that connects them, and we denote it by $d(i,j)$. If in a given network there does not exist any path that connects individuals i and j, we say that the distance between them is infinite, and $d(i,j) = \infty$. For example, in a star-shaped network any two different agents in the periphery are connected by a path of distance two. Since there is no other shorter path that connects these two peripheral agents, the distance among them in the network is equal to two. Finally, we say that a link among individuals i and j is a *bridge link* whenever these two individuals are of different types. Formally, the link ij is a bridge link if $\tau(i) \neq \tau(j)$. Bridge links are the ones that connect both communities.

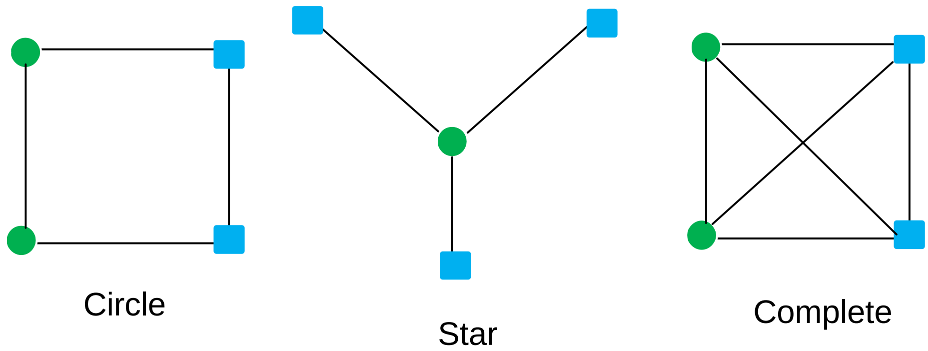


Figure 2. Circle, star and complete networks with four individuals.

2.2 Preferences

The utility function of each individual i, denoted by $u_i(g)$, depends on the network structure that connects all the population. It is given by

$$u_i(g) = \sum_j \delta^{d(i,j)} - \sum_{j \in \mathbb{N}_i(g)} c_{ij}(g) \tag{1}$$

where $0 \leq \delta < 1$ is the benefit from links, $d(i,j)$, the *geodesic distance* between individuals i and j, and $c_{ij} > 0$ is the cost for individual i of maintaining a direct link with j.

The utility function (1) has the general structure of the so-called *connections model*, introduced by Jackson and Wolinsky (1996). Links represent friendship relationships between individuals and involve some costs. A "friend of a friend" also results in some indirect benefits, although of a lesser value than the direct benefits that come from a "friend". The same is true of "friends of a friend of a friend," and so forth. The benefit deteriorates in the geodesic distance of the relationship. This is represented by a factor δ that lies between 0 and 1, which indicates the benefit from a direct relationship between i and j, and is raised to higher powers for more distant relationships. For instance, in the network described in Figure 3, individual 1 obtains a benefit of 2δ from the direct connections with individuals 2 and 3, an indirect benefit of δ^2 from the indirect connection with individual 4, and an indirect benefit of $2\delta^3$ from the indirect connection with individuals 5 and 6. Since $\delta < 1$, this leads to a lower benefit of an indirect connection than of a direct one.

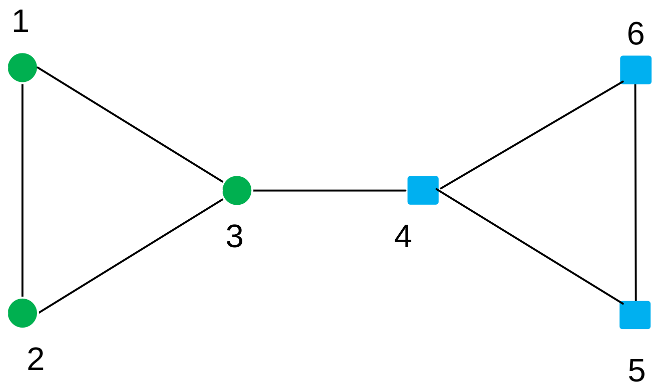


Figure 3. A bridge network.

However, individuals only pay costs $c_{ij} > 0$ for maintaining their *direct* relationships. This is where our model becomes very different from the standard connections model. To characterize linking costs, we need first to introduce one more concept. Given a network g, we define the *rate of exposure* of individual i to her own community $\tau(i)$ as:

$$e_i^{\tau(i)}(g) = \frac{n_i^{\tau(i)}(g)}{n_i(g) - 1}. \tag{2}$$

This ratio $e_i^{\tau(i)}(g)$ measures the fraction of *same-type friends* of individual i in network g since $n_i^{\tau(i)}(g)$ is the number of i's same-type friends in network g while $n_i(g)$ is the total number of i's friends in network g, independently of their type. The reason why we substract 1 in the denominator will become apparent in the next paragraphs.

We can now introduce the cost structure. Let c and C be strictly positive constants. We assume that:

$$c_{ij}(g) = \begin{cases} c & \text{if } \tau(i) = \tau(j) \\ c + e_i^{\tau(i)}(g)e_j^{\tau(j)}(g)\, C & \text{if } \tau(i) \neq \tau(j) \end{cases} \tag{3}$$

There are thus different costs, depending with whom a connection is made. Since $C > 0$ and the rate of exposure are non-negative, the main feature of this cost structure is that it is always more costly to form a friendship relationship with someone from the other community (the cost of which is $c + e_i^{\tau(i)}(g)e_j^{\tau(j)}(g)\, C$) than with someone from the same community (the cost of which is c). In particular, if an individual i of type $\tau(i)$ forms a friendship relationship with an individual j of type $\tau(j)$, with $\tau(i) \neq \tau(j)$ (i.e. inter-community friendship formation), then the cost is increasing in their respective rates of exposure to their own communities. If, for example, a green person has only green friends, then it will be difficult for her to interact with a blue person, especially if the latter has mostly blue friends. There are different cultures, norms and habits between communities so that frictions are higher the more different the people are. If we interpret "type" by "race" so that "blue" and "green" are replaced by "black" and "white", then (3) means that it is always easier for blacks to interact with other blacks and likewise for whites, and that the interracial relationships strongly depend on how "exposed" individuals are, i.e. how many same-race friends they have. These difficulties in interracial relationships can be due to language barriers[9] or more generally to different social norms and cultures.[10,11]

What we have in mind here is that individuals are born with a certain type τ (blue or green) that affects how easily they interact with other individuals. It is assumed that it is less costly to interact with someone of the same type than of a different type. So from this initial trait τ, there are natural gaps and differences between communities of types. But people make choices in terms of friendships. These choices can increase or decrease the original gap between individuals. If someone who is born blue chooses to only have blue friends, then it will be more difficult for her to interact with a green person. However, the more similar the friendship composition of two individuals of different types, the easier it is for them to interact. Observe that we allow friend choice to totally erase the initial cost gap between a blue type and a green type. Indeed, if at least one individual (i or j) has no friends of the same type (i.e. $e_i^{\tau(i)} = 0$ or $e_j^{\tau(j)} = 0$), then it is equally costly for them to interact with someone of the opposite type as with someone of the same type (i.e. the cost is c in both cases).[12]

[9] For example, the studies of Labov (1972), Baugh (1983), and Labov and Harris (1986) reveal that Black English of different metropolitan areas has converged, while it has been simultaneously diverging from Standard American English. This creates some costs in the interactions between blacks and whites.

[10] Camargo et al. (2010) show in a randomized experiment that whites who are randomly assigned black roommates have a significantly larger proportion of black friends in the future than white students who are randomly assigned white roommates. Ben-Ner et al. (2009) show in lab experiments that the distinction between in-group and out-group significantly affects economic and social behavior, for example, in forming working relationships.

[11] Lemanski (2007) documents an interesting experiment in post-apartheid urban South Africa by examining the lives of those already living in desegregated spaces. She studies the case a low-cost state-assisted housing project situated in the wealthy southern suburbs of Cape Town. In this social housing project, named Westlake village, colored and black African (alongside a handful of white and Indian) residents were awarded state housing in 1999 as replacement for their previous homes, which were demolished to make way for a mixed land-use development. She finds that different races are not only living peacefully in shared physical space but also actively mixing in social, economic and to a lesser extent political and cultural spaces. Furthermore, residents have largely overcome apartheid histories and geographies to develop new localized identities. This can be another indication that when people from different races or cultures interact with each other the costs of further interaction decreases.

[12] In Section 3.2, we investigate a different cost function where the intercommunity cost is *no* longer equal to the intracommunity even if one of the persons involved in a relationship has no friends of the same type.

The reason why we substract 1 in the denominator in the definition of the rate of exposure (see (2)) is because, when we compute the cost of a given bridge link between communities, we do not include this bridge link in the computation of the cost. What is relevant for the cost is the rate of exposure according to the rest of connections of each of the two individuals involved in the bridge link.

To illustrate our cost function (3), consider again the network described in Figure 3 and assume that individuals 1, 2, and 3 are greens (circle nodes) while individuals 4, 5, and 6 are blues (square nodes). Circle and square nodes represent green and blue individuals, respectively.Imagine that individuals 3 and 4 are not yet connected and individual 3 considers the possibility of creating a link with 4. In that case, the cost of connecting 3 (green) to 4 (blue) is:

$$c_{34}(g) = c + \frac{n_3^{\tau(3)}(g)}{n_3(g) - 1} \frac{n_4^{\tau(4)}(g)}{n_4(g) - 1} C = c + C$$

since $n_3^{\tau(3)}(g) = n_4^{\tau(4)}(g) = 2$ (number of same-type friends of 3 and 4, respectively) and $n_3(g) = n_4(g) = 3$ (total number of 3's and 4's friends independently of type, considering also the link between them),[13] which implies that $e_3^{\tau(3)}(g) = e_4^{\tau(4)}(g) = 1$.

If, for example, individual 4 also had a link with 2, the cost of connecting 3 (green) to 4 (blue) would be

$$c_{34}(g) = c + \frac{n_3^{\tau(3)}(g)}{n_3(g) - 1} \frac{n_4^{\tau(4)}(g)}{n_4(g) - 1} C = c + \frac{2}{3} C$$

since $e_3^{\tau(3)}(g) = 1$ but $e_4^{\tau(4)}(g) = 2/3$. It would be less costly for individual 3 (green) to befriend individual 4 (blue) in this situation because the latter has already a green friend.[14]

With the above notation, we want to highlight that costs, in particular inter-community costs, depend on the network structure. However, from now on, so as to minimize notational burden, we will not make the dependency of the rates of exposure and the linking costs on g explicit.

[13] Observe that, when individual 3 considers the possibility of creating a link with individual 4, individual 3 does not take into account the possible link between 3 and 4 when calculating the percentage of her own and 4's same-race friends.

[14] There are clearly other ways of defining the cost function than (3) since the latter may have some shortcomings. Consider, for example, a green agent (say agent 1) who has 2 links, one with a green agent and one with a blue agent and another green agent (say agent 2) with 99 links with green agents and one link with a blue agent. Using (2), it is easily verified that both agents 1 and 2 will have the same exposure rate equal to 1. This is true despite the fact that green agent 1 has 50% of her friends who are blue whereas green agent 2 only has 1% of her friends who are blue. We believe, however, that our definition is still reasonable, even in this example. Indeed, in both cases, when considering forming a link with a green agent $i = 1, 2$, what matters for the blue agent is that this green agent $i = 1, 2$ has a zero rate of exposure because all her friends are green and it does not matter whether the green agent i has only one friend ($i = 1$) or 99 friends ($i = 2$) who are green. What matters is really how "exposed" is the green agent to the blue community. In both cases, before the link between the blue and the green agent is formed, the green agent has zero exposure to the blue community, independently of how many green friends she has.

2.3 Network stability

In games played in a network, individuals' payoffs depend on the network structure. In our case, this dependency is established in expression (1), that encompasses both the benefits and costs attributed to an individual given her position in the network of relationships. Any equilibrium notion introduces some stability requirements. The notion of *pairwise-stability*, introduced by Jackson and Wolinsky (1996), provides a widely used solution concept in networked environments. Define by $g + ij$ the network g where the link ij has been added and by $g - ij$ the network g where the link ij has been removed.

Definition 1 *A network g is* ***pairwise stable*** *if and only if:*

(i) *for all $ij \in g$, $u_i(g) \geq u_i(g - ij)$ and $u_j(g) \geq u_j(g - ij)$*

(ii) *for all $ij \notin g$, if $u_i(g) < u_i(g + ij)$ then $u_j(g) > u_j(g + ij)$.*

In words, a network is pairwise-stable if (i) no player gains by cutting an existing link, and (ii) no two players not yet connected both gain by creating a direct link with each other. Pairwise-stability thus only checks for one-link deviations.[15] It requires that any mutually beneficial link be formed at equilibrium but does not allow for multi-link severance.

We will use this equilibrium concept throughout this paper. Thus, network g is an equilibrium network whenever it is pairwise stable.

3 Stable networks

3.1 Low intra-community costs

We start the analysis of stable networks with the case of low intra-community costs c. In particular, we start off assuming that $c < \delta - \delta^2$. If there were only one community (i.e. only one type of individuals), then the complete network would be the unique equilibrium network (as in the connections model of Jackson and Wolinsky, 1996). But, since we have two different communities and different cost structures, this is no longer true. Indeed, an individual of one type may decide to lower the exposure to her own community in order to become more attractive to the other one. This means, in particular, that $c < \delta - \delta^2$ cannot even guarantee that each community is fully intra-connected. Let us start by giving an example that illustrates this point. Consider the network depicted in Figure 4 where individuals 1, 2 and 3 are greens (circle nodes) while individuals 4, 5 and 6 are blues (square nodes). Let us show that, under $c < \delta - \delta^2$, the green individuals will *not* form a fully-connected society. Indeed, individual 3 will not want to form a link with 1 if and only if:

$$\delta - \delta^2 - c - \frac{C}{3} < 0 \qquad (4)$$

[15] This weak equilibrium concept is often interpreted as a necessary conditions for stronger stability concepts.

To understand this expression, observe that the benefit for 3 of forming a link with 1 is $\delta - \delta^2$ but the *net* cost is $c + C/3$ (c for the direct *intra-community* link with 1 and $C/3$ for maintaining the *inter-community* links with 4 and 5). Indeed, by creating the link 31, the rate of exposure of individual 3 increases from 1/2 (before the link 31) to 2/3 (after the link 31)[16] and thus the cost of maintaining the links to 4 and 5 increases from $2\,(c + C/2)$ to $2\,(c + 2C/3)$. This means that, even if $c < \delta - \delta^2$, the green community will not be fully intra-connected because, if (4) holds, 3 will not form a link with 1.[17]

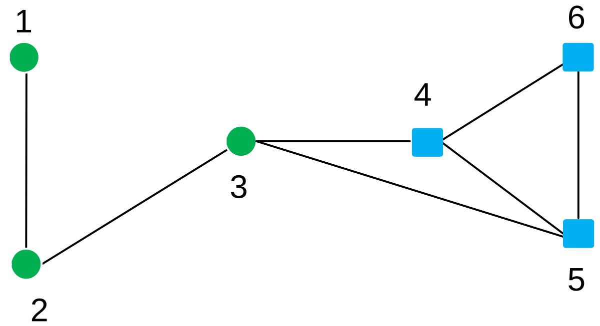


Figure 4. A pairwise-stable network

Let us now characterize the pairwise-stable equilibria for which each community is *fully intra-connected* (i.e. each individual is linked to all other individuals within the same community). We use the following definitions:

Definition 2

- *A network displays* ***complete integration*** *when both communities are completely connected (i.e. each community is fully intra-connected and both communities are fully inter-connected),*
- *A network displays* ***complete segregation*** *when both communities are isolated (i.e. each community is fully intra-connected but has no links at all with the other community),*
- *A network displays* ***partial integration*** *in any other case.*

We have the following result:[18]

[16] Observe that the rate of exposure of 4 and of 5 is not affected and is equal to 1.

[17] This is only true because 3 has some links with the blue community. Observe, however, that, if C is large enough, then the completely segregated network is an equilibrium network, where clearly the green community is fully connected.

[18] All proofs can be found in the Appendix.

Proposition 1 *Assume*

$$c < \delta - \delta^2 \tag{5}$$

(*i*) *The network such that the blue and the green communities are* ***completely integrated*** *is an equilibrium network if and only if*

$$C < \frac{(n-2)^2 (n-3)}{n^G (n^G - 1)^2} \left(\delta - \delta^2 - c\right) \tag{6}$$

(*ii*) *The network for which the blue and the green communities are* ***completely segregated*** *is an equilibrium network if and only if*

$$C > \delta + \left(n^G - 1\right) \delta^2 - c \tag{7}$$

To interpret these results, it is useful to think about two different effects. The first effect, which we refer to as the *connections effect*, expresses the role of the direct and indirect gains and losses of forming or severing a link. This first effect, which is also present in the connections model of Jackson and Wolinsky (1996), means that direct connections give higher utility than indirect connections. There is, however, a second effect, that we refer to as the *exposure effect*, which is new. This effect is due to the fact that the formation of a new link affects the exposure rates of the individuals involved in it. Indeed, if the new link is between two individuals from different communities (the same community), then the rate of exposure of each of these individuals to their own community is going to decrease (increase) and thus their inter-community costs will decrease (increase). This indirect exposure effect is positive (negative).

As a result, the completely integrated network is going to be stable if the sum of the connections and the exposure effects for any link is positive. Consider an inter-community link. The connections effect is ambiguous because the cost of keeping the link for each individual is strictly larger than c since their rates of exposure to their own communities are strictly positive. However, severing such a link has an strong and negative *exposure effect* since it increases both their rate of exposure and the inter-community costs with the rest of their friends. Some algebra shows that this second (exposure) effect always dominates the connections effect and, hence, nobody has an incentive to sever a link. The case of an intra-community link is less clear. In such a case, the connections effect can be positive if $\delta - \delta^2 - c > 0$. This would imply that for two individuals from the same community the benefits of a direct connection compared to an indirect connection of distance two always outweigh the costs of forming such a link. However, keeping such link has a negative exposure effect: it increases their respective rates of exposure to their own communities, and therefore the costs of their inter-community links become larger. If C is sufficiently low, then the negative exposure effect dominates the positive connections effect and we end up with a stable completely integrated network (see (6)).

The completely segregated network arises when the connections effect of an inter-community link is negative. Condition (7) is precisely the mathematical formulation of this negative effect. Note that, in this case, there are no exposure effects to consider since we start from a situation where there are no inter-community links. In that case, C now has to be high enough for this network to be an equilibrium.

Denote $\Delta \equiv \delta - \delta^2 - c$. The following proposition characterizes some partially integrated equilibrium networks, and bring into the picture a third important component in the stability of a network geometry:

Proposition 2 *Assume (5).*

(i) If

$$\max\left\{\frac{\Delta\, n^B}{(n^B-2)}, \frac{n^B\left[\Delta + \left(n^B-1\right)\left(\delta^2-\delta^3\right)\right]}{(n^B-1)}, \Delta + \left(n^G-1\right)\left(\delta^2-\delta^3\right)\right\} < C < \Delta + n^B\delta^2 \tag{8}$$

holds, then the network where both communities are fully intra-connected and where there is only one bridge link is an equilibrium network (Figure 5).

(ii) If

$$\frac{\Delta\, n^G n^B\left(\delta - \delta^2 - c\right)}{(n^G-1)(n^B-1)-n^G} < C < \delta - \delta^3 - c \tag{9}$$

holds, then the network where both communities are fully intra-connected and each blue individual has one, and only one, bridge link and where each green individual has at most one bridge link is an equilibrium network (Figure 6).

(iii) If

$$\frac{n^G(n-2)\,\Delta}{(n^G-1)(n-2)-(n^B-1)} < C < (n-2)\min\left\{\frac{(n-3)\,\Delta}{(n^G-1)^2}, \frac{\Delta + \left(n^B-1\right)\left(\delta^2-\delta^3\right)}{n^B-1}\right\} \tag{10}$$

holds, then the network in which both communities are fully intra-connected and only one blue agent connects with the green community by linking to all green individuals is an equilibrium (Figure 7).

Observe that the conditions in (*i*), (*ii*) and (*iii*) given in Proposition 2 are *not* mutually exclusive. In these equilibrium configurations some integration between greens (circle nodes) and blues (square nodes) is taking place. The following figures provide a graphical representation.

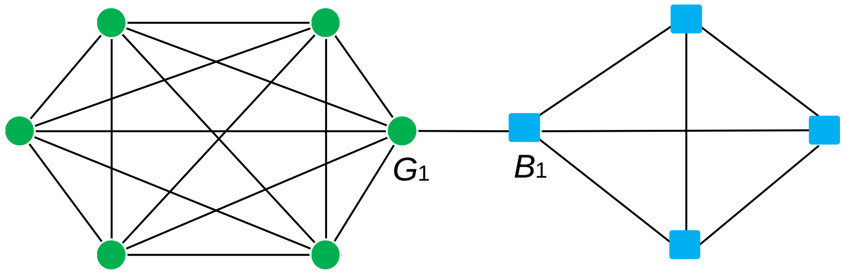


Figure 5. Equilibrium network when condition (8) holds.

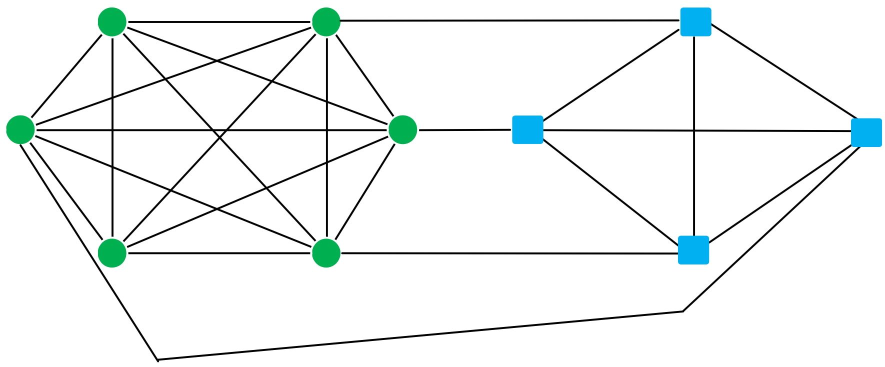


Figure 6. Equilibrium network when condition (9) holds.

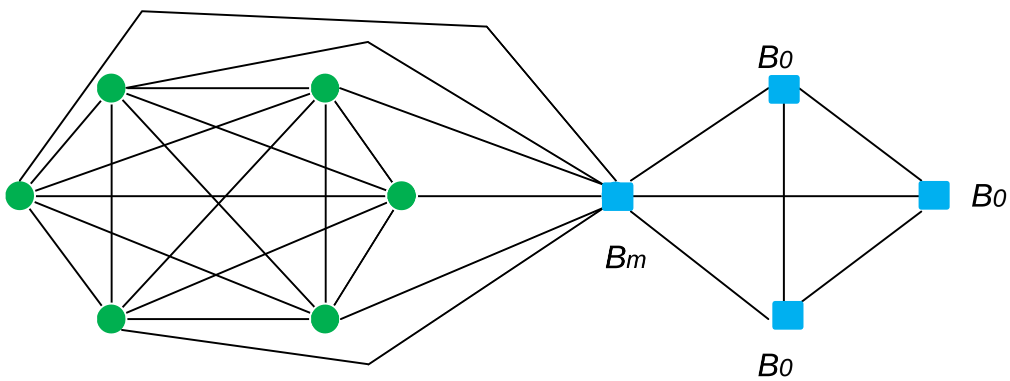


Figure 7. Equilibrium network when condition (10) holds.

As before, both connections and exposure effects are present in explaining the results of Proposition 2. There is, however, a third component that becomes relevant here: the requirement of *mutual consent* for the link to be formed. When the inter-community costs of forming a link are relatively large, i.e. C is high, the network in Figure 5 is pairwise stable because the connections effect for the agents involved in the only bridge link between communities is positive while the connections effect of any other inter-community link is negative for at least one of the two sides of each of these potential links.[19]

When C decreases slightly, individuals from different communities may now want to create one of these missing links. This is illustrated in Figure 6. While the direct benefits of such a new connection have not changed, the costs are now reduced and, as a result, the sign of the connections effect of such a new link is reverse. In both networks described in Figures 5 and 6, the exposure effects play no role since each of the agents is involved in at most one link and the cost of this link is kept constant when there are changes in the connections within the community (these intra-community links do not change the rate of exposure of individuals, which remain maximal and equal to 1, according to the definition of the rate of exposure given in (2); see Lemma 1 in the Appendix).

The logic behind the stability of the network displayed in Figure 7 is different since it strongly relies on the exposure effect. The B_m blue individual invests in a large number of inter-community links in order to decrease her own rate of exposure enough and thus to decrease her own cost of each of these connections. This, in turn, makes it cheaper for each green individual to connect to her and win direct access to the blue community.

To understand our results, let us summarize the three main forces at work:

(1) Individuals want to form connections to obtain direct and indirect benefits. In a disperse network, connecting to a member of a different community usually gives access to many opportunities. This is the *connections effect*.

(2) Because links are costly, individuals become more attractive the more they are friends with individuals from the other community and hence can form new links more easily with the other community. This is the *exposure effect*.

(3) There is a *coordination problem* because the creation of a link needs the consent of both individuals. Condition (*ii*) in Definition 1 of pairwise stability highlights this *mutual consent effect*.

Equilibrium networks are those that correctly balance these three forces at the individual level. The equilibrium networks characterized in Proposition 1 and 2 provide some understanding on how these three effects interact with each other. Contrary to the literature on segregation (e.g. Schelling, 1971; Benabou, 1993), it is important to observe that, here, both the *individual location* and the *structure of the network* are crucial to understand the equilibrium outcomes. Indeed, not

[19] The two individuals involved in this bridge link enjoy a singular position in the network. Some literature in sociology has highlighted the importance of these type of links in terms of social capital: it is important that bridges exist between communities. Indeed, social capital is created by a network in which people can broker connections between otherwise disconnected segments (Granovetter, 1973, 1974; Burt, 1992). We can say that the people who are bridging two communities are sitting in a *structural hole* of the network. A structural hole exists when there is only a weak connection between two clusters of densely connected people (Burt, 1992; Goyal and Vega-Redondo, 2007).

only benefits but costs are affected by an individual's location and the structure of the network. For example, two identical blue individuals who have different positions in the network may have different incentives to form a link with a green person so that, in equilibrium, only one of them will find it beneficial to form a bridge link.

The next result shows under which condition it is even possible that all agents in an economy are such that none of them has a friend with someone from her own community.

Proposition 3 *If*

$$\delta - \delta^2 < c < \delta - \delta^3 \tag{11}$$

the bipartite network in which all green agents are connected to all blue agents, and all blue agents are connected to all green agents is an equilibrium network (Figure 8).

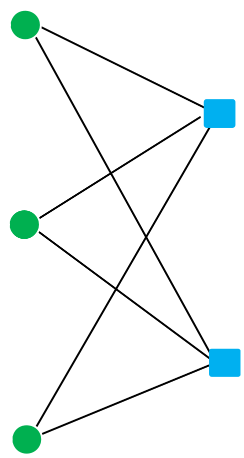


Figure 8. Bipartite network with $n^G = 3$ and $n^B = 2$.

Proposition 3 shows that the bipartite network is an equilibrium network if (11) holds. Observe first that, contrary to the previous result in Proposition 2 case (*iii*) (see condition (10) for the network in Figure 7), here condition (11) is not a function of the inter-community cost C. This is because *both* the green (circle node) and the blue (square node) agents have "extreme" behavior in the sense that they are only friends with individuals from the other community. As a result, their exposure rate is always zero and when someone wants to deviate from the bipartite network (by either creating or deleting a link), the inter-community cost is equal to the intra-community cost, which is c. Take for example the case when a green agent wants to create a link with another green agent. In that case, her exposure rate increases from 0 to $1/n^B$ but the exposure rate of all the blue agents she is connected to is still 0 (since they have no link with each other) and thus the inter-community cost of the green agent is still c and does not depend on C. The same type of reasoning applies for the severance of a link. This result is due to the fact that both exposure rates matter in inter-community costs, highlighting the role of mutual consent in the friendship decision. Observe also that condition (11), i.e. $\delta - \delta^2 < c < \delta - \delta^3$, is not compatible with condition (5), i.e. $c < \delta - \delta^2$, which was a necessary condition for communities to be fully intra-connected.

3.2 More general cost function

In the previous section, we found that bipartite networks or other networks with "extreme" behaviors were pairwise stable (see Proposition 2 (*iii*) and Proposition 3) because there were no costs of becoming "green" for a blue person. For example, in the equilibrium bipartite network described in Figure 8, a blue agent "becomes" a green individual for the other green agents since the cost of interacting with her is just c, as for all the other intra-community costs. This is due to our assumption on the cost function, which stipulates that the *inter-community* cost is equal to the *intra-community* cost as soon as one of the persons involved in the relationship has no friends of the same type. In the present section, we relax this assumption and assume instead the following *inter-community* cost function for $\tau(i) \neq \tau(j)$:

$$c_{ij} = c + \left(k + e_i^{\tau(i)} e_j^{\tau(j)}\right) C \qquad (12)$$

where $0 < k < 1$ (we still assume that $c_{ij} = c$ if $\tau(i) = \tau(j)$). With this new inter-community cost function, a blue person can never become totally "green" for other green individuals because even if she has no blue friends, i.e. $e_i^{\tau(i)} = 0$, the cost of interacting with greens is $c + k\,C$, which is strictly greater than c, the cost for a green of interacting with other greens.

Proposition 4 *Consider the inter-community cost function given by (12) and assume that $c < \delta - \delta^2$.*

(i) If

$$C < \frac{\left(n^G + 1\right)\left(\delta - \delta^2 - c\right)}{n^G} \qquad (13)$$

holds, then any equilibrium network is such that each community is ***fully intra-connected****. In particular, a bipartite network (such as the one described in Figure 8) can never be an equilibrium.*

(ii) Furthermore, if

$$C < \left(\delta - \delta^2 - c\right) \min\left\{ \frac{(n-2)^2\,(n-3)}{n^G\left(n^G - 1\right)^2}, \frac{1}{k + \frac{(n^B-1)(n^G-1)(n^G-2)}{(n-3)(n-2)^2}} \right\} \qquad (14)$$

holds, then the network for which the blue and green communities are ***totally integrated*** *is an equilibrium network.*

(iii) If

$$C > \frac{\delta + \left(n^G - 1\right)\delta^2 - c}{1 + k} \tag{15}$$

holds, then the network for which the blue and green communities are ***completely segregated*** *is an equilibrium network.*

When the inter-community cost function is given by (12), then each community forms a complete network if C is not too large. In that case, no bipartite network can emerge. This is because now nobody can become "like" someone from the other type and, therefore, the attractiveness of only having friends from the other community is much lower. Interestingly, when k and/or C are not too large, then each individual will have links with all individuals (including those from the other community). The two communities are totally integrated. On the contrary, if k and/or C are large enough, then only links with her own community prevail and the two communities are completely segregated. Indeed, once the network is totally integrated, then nobody wants to delete a link because the gain is too low compared to the costs (this is because k is low enough). When the network is completely segregated, then because C is high enough, no individual wants to form a link with someone from the other community.

More generally, observe that, with the inter-community cost function given by (12), green individuals never become blues for blue individuals and vice versa. However, the change is very small because, compared to (3), it just adds a constant k. Let us illustrate this point by showing under which condition the bipartite network described in Figure 8 is an equilibrium network. With the cost function (3), condition (11), i.e. $\delta - \delta^2 < c < \delta - \delta^3$, guaranteed that this was an equilibrium network. With the cost function (12), it is easily verified that the condition is now: $\delta - \delta^2 < c < \delta - \delta^3 - kC$. It is very close but the main difference is the fact that C (and k) now appear in the condition. In particular, C (and k) has to be low enough, otherwise inter-community links will not be formed. This, however, does not affect the fact that exposure rates are still equal to zero and that deviating from the bipartite network by, for example, creating a link with your own community, does not affect the cost of maintaining the inter-community links.

3.3 Higher socialization costs

Let us now go back to our original cost function (3) and consider the case when $c > \delta - \delta^2$ so that it becomes more expensive to form links with individuals from the same community. The cost structure is as in the benchmark model and given by (3). In that range of parameters (i.e. $\delta - \delta^2 < c < \delta$), Jackson and Wolinsky (1996) have shown that, for each community, a star encompassing all individuals is always a pairwise stable network.[20] We thus focus on communities that have a star-shaped form. Of course, since we are dealing with a different cost structure, it is

[20] Observe that it is not necessarily the unique pairwise stable graph.

not necessarily true that this result remains valid. However, we are going to present a family of equilibrium networks in which intra-group structure always forms a star network.

Proposition 5 *Assume that*

$$\delta - \delta^2 < c < \delta \tag{16}$$

(i) If

$$C > \delta + \left(n^G - 1\right)\delta^2 - c \tag{17}$$

*then the network with two disconnected star-shaped communities is a pairwise equilibrium network (**complete segregation;** see Figure 9(i)).*

(ii) If

$$C > \delta - \delta^3 - c \tag{18}$$

*then the network where the star-shaped communities are connected through their central agents is a pairwise equilibrium network (**partial integration;** see Figure 9(ii)).*

(iii) If

$$c > \delta - \frac{2}{5}\delta^4 + \frac{\delta^2}{5} - \frac{4}{5}\delta^3$$

and

$$\max\left\{\delta - 2\delta^4 + \delta^2 - c, 2\left(\delta - \delta^3 + \delta^2 - \delta^4 - c\right)\right\} < C < 4\left(c - \delta + \delta^3\right) \tag{19}$$

*then the network where each peripheral agent in the star-shaped community has one bridge link with the other peripheral agent whereas stars have no bridge links is a pairwise equilibrium network (**partial integration**; see Figure 9(iii)).*

(iv) If

$$C < \delta - \delta^3 - c \tag{20}$$

*then the network where the centers in both star-shaped communities are connected to each other and all peripheral agents from both communities are connected to each other is a pairwise equilibrium network (**partial integration;** see Figure 9(iv)).*

Figure 9 displays the different cases of Proposition 5 for $n^B = n^G = 3$.

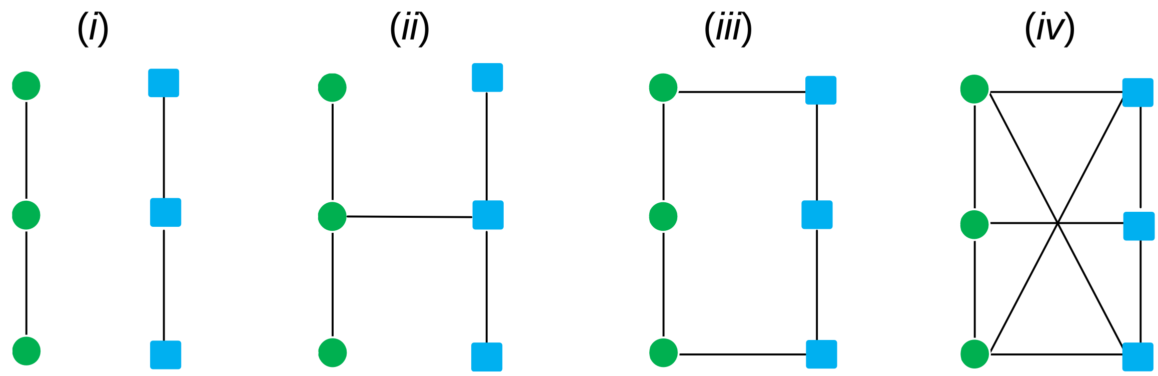


Figure 9. Different equilibrium networks when $\delta - \delta^2 < c < \delta$.

These results are quite intuitive and show how a reduction in C leads to more bridge links and more interactions between communities. Let us explain, for example, why in case (iv), some blue agents (square nodes) are mostly friends with other blue agents while other blue agent are mostly friends with green agents (circle nodes). Indeed, in case (iv), each peripheral blue (green) individual has one blue (green) friend (the central agent) and $n^G - 1$ $(n^B - 1)$ green (blue) friends so that their common same-type friend percentage is $e_i^{\tau(i)} = 1/(n^{\tau(i)})$. This is quite small, especially when the size of the population of each community is large. As a result, each blue (green) peripheral individual displays a high taste for other-type friends, which makes them very attractive. On the contrary, the blue (green) central agent has one green (blue) friend and $n^B - 1$ $(n^G - 1)$ blue (green) friends so that $e_i^{\tau(i)} = (n^{\tau(i)} - 1)/n^{\tau(i)}$. This percentage is very close to 1, which make this central agent less attractive for people from the other community.

It is now easy to understand why we have these extreme behaviors. Let us focus on blues. First, peripheral blues do not want to connect to each other because the cost is too high compared to the benefit since $c < \delta - \delta^2$ (they are at distance 2 from each other). Second, peripheral blues do not want to sever a link with one of the $n^G - 1$ peripheral greens because the latter are all very attractive. Finally, peripheral blues do not want to create a link with a central green person because she is not very attractive due to her high inter-community costs and they can reach him from a peripheral green (distance 2) and obtains δ^2. This is why peripheral blues have most of their friends who are greens. It is also easy to understand why a blue central individual has most of her friends who are blues. This is due to the fact that she is not attractive (because of her high exposure to her own community) to the peripheral greens. It is important to observe that this result is *not* due to the size of the communities. It is easy to verify that it still holds if $n^B = n^G = n/2$. More generally, we can see here that there are *reinforcing effects* because once someone from one community is connected to someone from the other community, she becomes more attractive to people from the other community because she costs less in the sense that she is less isolated.

4 Social welfare: Integration versus segregation

We now consider some welfare implications of our model. We have previously focused on how decentralized linking decisions can lead to different social network structures. In particular, our model naturally leads to multiple equilibria. Our analysis in Proposition 1, for example, shows that there is a range of parameters in which two extreme outcomes, the complete network (in which all pair of agents, whatever their types, are connected) and a segregated network (in which only connections within communities are established) are both stable networks. The former represents a situation of *social integration* while the latter represents *social segregation*. In terms of efficiency considerations, one may wonder which of the two outcomes is better from a social viewpoint. Here, we shed some light on this issue by showing the most important source of inefficiencies in our model.

We consider a utilitarian perspective, where social welfare is measured by the unweighted sum of individual utilities. Thus, a network g is *socially preferable* to another network g' whenever the sum of individual utilities in g is higher than the sum of individual utilities in g', i.e. $\sum_i u_i(g) > \sum_i u_i(g')$.

The following result compares the social welfare of segregated and integrated networks, and states which one is socially preferable.

Proposition 6 *Assume $c < \delta - \delta^2$ and (7). If*

$$n^B \left(n^G - 1\right)\left(n^B - 1\right) \leq (n-1)^2 \tag{21}$$

holds, then there exists a threshold $\widetilde{C}$ such that for $C \leq \widetilde{C}$, integration is efficient whereas when $C \geq \widetilde{C}$, segregation is efficient.

This result suggests that, depending on the size of relative social groups, we can not plead for integrated or segregated socialization patterns *a priori*. The possible inefficiency of the integrated network comes from the fact that, for any individual, it is costly to keep all her inter-community links. Indeed, stability means that it is suboptimal for her to severe one intra-community link to increase the exposure effect, and therefore she is paying a cost that is proportional to C. When C increases, these costs may overcome the benefits derived from connecting to the other community. Note that this does not contradict stability: when all the rest of her community is connected to the other community, it is optimal for her to also connect to the other community. This is because the exposure rate of any member of the other community, and therefore the cost of directly connecting with each of them, is low precisely because of these inter-community connections with the rest of the group. Yet, from a collective point of view, each community would be better off in isolation because the aggregate socialization costs are too large when C lies above the threshold $\widetilde{C}$. As a result, the effects of exposure on costs can explain the possible inefficiency of interactions.[21]

[21] Throughout the discussion on the welfare issues, we assume that utilities are exogenously fixed across time and we disregard any possible positive aggregate effects of communication across different types.

This result allows us to link the cost mechanism we explore in this paper with some policy issues. In particular, we can extract some preliminary conclusions on the possible (in)effectiveness of policies that can favor socialization and thus interaction between different communities. Policies that diminish intra-community socialization costs are not necessarily going to induce more desirable network structures. For example, activities outside the classroom for adolescents or cultural activities at the neighborhood level can favor integrated patterns since they may facilitate interactions among individuals with different identities but the outcome is not going to be socially efficient unless these policies sufficiently decrease the cost of interactions. While the integrated network can be sustained in equilibrium, this equilibrium can be socially undesirable because individuals are exerting an excessive cost to keep their connections with the other community.

5 Social norms

We would like to extend our model to discuss how social norms and, in particular, social punishments for deviating from the social behavior of the rest of the community, can also influence friendship behaviors. For that, we modify the cost of socialization choices by taking into account *social norms*. There are studies that illustrate the importance of social sanctions and social norms in ethnic groups (see Akerlof, 1997, and references therein). Anson (1985) relates the story of Eddie Perry, an African-American youth from Harlem, who graduated with honors from Phillips Exeter Academy and won a full four-year fellowship to Stanford. A close mentor of Eddie explained the psychological tension of coming back home in his own neighborhood: "This kid couldn't even play basketball. They ridiculed him for that, they ridiculed him for going away to school, they ridiculed him for turning white. I know because he told me they did." (Anson, 1985, p. 205). In his autobiographical essay, Rodriguez (1982) told us about his own story as a Mexican-American from Sacramento who went to college and for whom English became his dominant language. His (extended) family considered him increasingly alien and as he put it: "*Pocho*, they called me. Sometimes, playfully, teasingly, using the tender diminutive −*mi pochito*. Sometimes not so playfully, mockingly, *Pocho* (Rodriguez (1982, p. 29).[22] These two stories of a black person labeled a white man by his black neighbors and an Hispanic labeled a "gringo" by his extended family are strikingly similar and illustrate the idea of social sanctions and social norms imposed by their own communities.[23]

In what follows, we propose a simple way of incorporating these forms of social norms and sanctions in our model. The two examples mentioned in the previous paragraph share the same characteristics: an individual of a given community is punished because she is deviating from the

[22] As Akerlof (1987) noted it, "a Spanish dictionary defines the word 'pocho' as an adjective meaning 'colorless' or 'bland'. As a noun it means the Mexican-American who, in becoming an American, forgets his native society.

[23] See also Stack (1976) for an interesting story of social sanctions/norms imposed by two sisters on their third sister who became middle class. Stack explained how the social distance between them increased, especially clear in the mutual care of their respective children.

social norms imposed by her community. In other words, social punishments increase when there are strong differences between own exposure rate and the exposure rate of own friends. We can formalize this idea using a *social sanction function* that depends on these rates of exposure. The social sanction individual i receives from her own community is a function of her rate of exposure, e_i, and the average rate of exposure of her friends from her own community, which we denote by $\bar{e}_i$ and is given by:

$$\overline{e}_i(g) = \frac{\sum_{j \in N_i^{\tau(i)}(g)} e_j^{\tau(j)}(g)}{n_i^{\tau(i)}(g)}.$$

To understand this formula, consider again the network described in Figure 7 and let us calculate the average rate of exposure of B_m's friends. We obtain:

$$\overline{e}_{B_m}(g) = \frac{3e_{B_0}^B(g)}{3} = e_{B_0}^B(g) = 1$$

We denote by $s(e_i, \overline{e}_i)$, the social sanction imposed to i. The utility function of individual i is then defined as:

$$u_i(g) = \sum_j \delta^{d(i,j)} - \sum_{j \in N_i(g)} c_{ij} - s(e_i, \overline{e}_i) \tag{22}$$

The utility of individual i includes now three different components: the benefits derived from direct and indirect connections, $\sum_j \delta^{d(i,j)}$, the total cost of forming direct links with both communities, $\sum_{j \in N_i(g)} c_{ij}$, and the social sanction imposed on individual i by her community, $s(e_i, \overline{e}_i)$.

According to our interpretation, there are several properties this social sanction function should satisfy:

(i) The social sanction is positive, i.e. $s(e_i, \overline{e}_i) > 0$, only when $0 < e_i < \overline{e}_i$ because, in that case, individual i spends more time with the other community than the average of her friends from her own community;

(ii) The social sanction is equal to zero, i.e. $s(e_i, \overline{e}_i) = 0$, if $e_i \geq \overline{e}_i$ or $e_i = 0$, that is when she spends either more time with her community than the average of her friends or no time at all (in which case, no sanction is possible);

(iii) In the case when $0 < e_i < \overline{e}_i$, the social sanction imposed on individual i is higher the larger is the difference between e_i and $\overline{e}_i$, i.e.

$$\frac{\partial s(e_i, \overline{e}_i)}{\partial e_i} \leq 0$$

(*iv*) Finally, when $0 < e_i < \overline{e}_i$, the effect of a decrease in individual i's rate of exposure is stronger when the average rate of exposure of the peer group of individual i in her own community is larger, that is:

$$\frac{\partial^2 s\left(e_i, \overline{e}_i\right)}{\partial \overline{e}_i \partial e_i} \leq 0$$

To fix ideas, consider the following social sanction function:

$$s\left(e_i, \overline{e}_i\right) = \left(e_i - \overline{e}_i\right)^2 \mathbf{1}_{\{0 < e_i < \bar{e}_i\}}$$

where $\mathbf{1}_{\{0 < e_i < \overline{e}_i\}}$ denotes the indicator function on the set $\{0 < e_i < \bar{e}_i\}$. This function satisfies the four above properties since: (*i*) $s\left(e_i, \overline{e}_i\right) = \left(e_i - \overline{e}_i\right)^2 > 0$ when $e_i < \overline{e}_i$; (*ii*) $s\left(e_i, \overline{e}_i\right) = 0$ when $e_i \geq \overline{e}_i$ or $e_i = 0$; (*iii*) and (*iv*) when $e_i < \overline{e}_i$, we have:[24]

$$\frac{\partial s\left(e_i, \overline{e}_i\right)}{\partial e_i} = 2(e_i - \overline{e}_i) < 0 \text{ and } \frac{\partial^2 s}{\partial \overline{e}_i \partial e_i} = -2 < 0$$

In this new scenario, the social sanction function amplifies the exposure effects since the social sanction adds an implicit cost for individual i in case her exposure to the other community is higher than the exposure of her peers. The direct benefits and cost of a given link are unaffected compared to the initial formulation of the model, which means that the *connections effects* in the two models coincide. To understand the consequences of this new term in the utility function, let us analyze some of the networks we studied before.

Consider, for example, the integrated and segregated networks described in Proposition 1. The social sanction function $s\left(e_i, \overline{e}_i\right)$ facilitates the stability of segregated networks. Indeed, when considering creating a link with the other community, for each individual, independently of one's type, the connections effect is unaffected because the externalities and the direct cost of building a link are the same as in the initial model. The exposure effect is, however, magnified because creating a bridge link will decrease the exposure rate of each individual involved in the link, which will be below that of the rest of the community. As a result, because of the social sanction, the incentives to create an inter-community link are lowered and the stability of segregated networks is preserved.

In some sense, Eddie Perry and Richard Rodriguez (mentioned above) have both chosen to have a very low exposure to their own community (i.e. low e_i) and are paying a very high price for it when interacting with people from their community of origin.

[24] Economists have modelled *conformity* in a similar way by adding a term $-(e_i - \overline{e}_i)^2$ to the utility function, where e_i and $\overline{e}_i$ is the effort of individual i and the average effort of i's peers. In that case, each individual i loses utility $(e_i - \overline{e}_i)^2$ from failing to conform to her peers (see, among others, Akerlof, 1980, 1997; Fershtman and Weiss, 1998; Patacchini and Zenou, 2012; Liu et al., 2014). Our formulation is slightly different since the social cost is only paid when $e_i < \overline{e}_i$.

6 Conclusion

We analyze a network formation model where agents belong to different communities. Both individual benefits and costs depend on direct as well as indirect connections. Two individuals from the same community always face a low linking cost while the cost of forming a relationship between two individuals from different communities diminishes with the rate of exposure of each of them to the other community. When intra-community linking costs are low, we show that two communities may integrated or segregated depending on the inter-community costs. We also show, that, in several equilibrium configurations, bridge links (i.e. links that connect both communities) prevail. We also find that socialization among the same type of agents can be weak even if the cost of maintaining links within one's own type is very low. Our model also suggests that policies aiming at reducing segregation are socially desirable only if they reduce the within-community cost differential by a sufficiently large amount.

In what follows, we suggest two avenues for future research that seem particularly promising.

First, from a more technical perspective, it would also be worth studying possible refinements of our equilibrium concept that could help providing more precise results and a more exhaustive characterization of the set of equilibrium networks. This is going to increase the already important combinatorial complexity in the analysis, which already deprives us from obtaining a full characterization of pairwise stable networks.

Second, we have not deepened other important consequence of network structure, such as segregation and inequality. Echenique and Fryer (2007) has introduced a new measure of individual segregation rooted at the social network level. This measure could be used in our setup to analyze the segregation patterns emerging from decentralized network formation. Kets et al. (2011) have also proposed an interesting model exploring how the structure of a social network constrains the level of inequality that can be sustained among its members. In their model, what influences inequality is the ability of players to form viable coalitions given an exogenous social network. It would be interesting to relate network formation and segregation considerations to these relevant issues.

References

[1] Anson, R.S. (1985), *Best Intentions: The Education and Killing of Edmund Perry*, New York: Random House.

[2] Benabou, R. (1993), "Workings of a city: Location, education, and production," *Quarterly Journal of Economics* 108, 619-652.

[3] Bramoullé, Y., Currarini, S., M.O. Jackson, Pin, P. and B.W. Rogers (2012), "Homophily and long-run integration in social networks," *Journal of Economic Theory* 147, 1754-1786.

[4] Braugh, J. (1983), *Black Street Speech: Its History, Structure and Survival*, Austin: University of Texas Press.

[5] Cabrales, A., Calvó-Armengol, A. and Y. Zenou (2011), "Social interactions and spillovers," *Games and Economic Behavior* 72, 339-360.

[6] Camargo, B., Stinebrickner, R. and T. Stinebrickner (2010), "Interracial friendships in college," NBER Working paper No. 15970.

[7] Coleman, J.S. (1958), "Relational analysis: the study of social organizations with survey methods," *Human Organization* 17, 28-36.

[8] Currarini, S., Jackson, M.O., and P. Pin (2009), "An economic model of friendship: Homophily, minorities, and segregation," *Econometrica* 77, 1003-1045.

[9] Currarini, S., Jackson, M.O., and P. Pin (2010), "Identifying the roles of race-based choice and chance in high school friendship network formation," *Proceedings of the National Academy of Sciences of the USA* 107, 4857-4861.

[10] de Martí, J. and Y. Zenou (2011), "Social networks," In: I. Jarvie and J. Zamora-Bonilla (Eds.), *Handbook of Philosophy of Social Science*, London: SAGE Publications, pp. 339-361.

[11] Echenique, F. and R.D. Fryer, Jr (2007), "A measure of segregation based on social interactions," *Quarterly Journal of Economics* 122, 441-485.

[12] Fershtman, C. and Y. Weiss (1998), "Social rewards, externalities and stable preferences," *Journal of Public Economics* 70, 53-73.

[13] Gallo, E. (2012), "Small world networks with segregation patterns and brokers," *Review of Network Economics* 11, Article 7.

[14] Goffette-Nagot, F., Grauwin, S. and P. Jensen (2012), "Dynamic models of residential segregation: An analytical solution," *Journal of Public Economics* 96, 124-141.

[15] Golub, B. and Jackson M. (2012), "How homophily affects the speed of learning and best-response dynamics," *Quarterly Journal of Economics* 127, 1287-1338.

[16] Goyal, S. (2007), *Connections: An Introduction to the Economics of Networks*, Princeton: Princeton University Press.

[17] Goyal, S. and F. Vega-Redondo (2007), "Structural holes in social networks," *Journal of Economic Theory* 137,460-492.

[18] Granovetter, M.S. (1973), "The strength of weak ties," *American Journal of Sociology* 78, 1360-1380.

[19] Granovetter, M.S. (1974), *Getting a Job: A Study of Contacts and Careers*, Cambridge, MA: Harvard University Press.

[20] Jackson, M.O. (2008), *Social and Economic Networks*, Princeton: Princeton University Press.

[21] Jackson, M.O. and B.W. Rogers (2005), "The economics of small worlds," *Journal of the European Economic Association* 3, 617-627.

[22] Jackson, M.O. and B.W. Rogers (2007), "Meeting strangers and friends of friends: How random are social networks?" *American Economic Review* 97, 890-915.

[23] Jackson, M.O and A. Wolinsky (1996), "A strategic model of social and economic networks," *Journal of Economic Theory* 71, 44-74.

[24] Jackson, M.O. and Y. Zenou (2015), "Games on networks", In: P. Young and S. Zamir (Eds.), *Handbook of Game Theory, Vol. 4*, Amsterdam: Elsevier Publisher, pp. 91-157.

[25] Johnson, C. and R.P. Gilles (2000), "Spatial social networks," *Review of Economic Design* 5, 273-299.

[26] Kets, W., Iyengar, G., Sethi, R. and S. Bowles (2011), "Inequality and network structure," *Games and Economic Behavior* 73, 215-226.

[27] Labov, W. (1972), *Language in the Inner City: Studies in the Black English Vernacular*, Philadelphia: University of Pennsylvania Press.

[28] Labov, W. and W.A. Harris (1986), "De facto segregation of black and white vernaculars, in D. Sankoff (Ed.), *Current Issues in Linguistic Theory 53: Diversity and Diachrony*, Amsterdam: John Benjamins Publishing, pp. 1-24.

[29] Lang, K. (2007), *Poverty and Discrimination*, Princeton: Princeton University Press.

[30] Lemanski, C. (2007), "The impact of residential desegregation on social integration: Evidence from a South African neighbourhood," *Geoforum* 37, 417-435.

[31] Liu, X., Patacchini, E. and Y. Zenou (2014), "Endogenous peer effects: Local aggregate or local average?" *Journal of Economic Behavior and Organization* 103, 39-59.

[32] Marmaros, D. and B. Sacerdote (2006), "How do friendships form?" *Quarterly Journal of Economics* 121, 79-119.

[33] Mayer, A. and S.L. Puller (2008), "The old boy (and girl) network: Social network formation on university campuses," *Journal of Public Economics* 92, 329-347.

[34] McPherson, M., Smith-Lovin, L. and J.M. Cook (2001), "Birds of a feather: Homophily in social networks," *Annual Review of Sociology* 27, 415-444.

[35] Mele, A. (2010), "A structural model of segregation in social networks," Unpublished manuscript, University of Illinois at Urbana-Champaign.

[36] Myerson, R.B. (1991), *Game Theory: Analysis of Conflict*, Harvard: Harvard University Press.

[37] Mobius, M. (2007), "The Formation of ghettos as a local interaction phenomenon," Unpublished manuscript, Harvard University.

[38] Patacchini, E., Picard, P.M. and Y. Zenou (2015), "Urban social structure, social capital and spatial proximity", CEPR Discussion Paper No. 10501.

[39] Patacchini, E. and Y. Zenou (2012), "Juvenile delinquency and conformism," *Journal of Law, Economics, and Organization* 28, 1-31.

[40] Patacchini, E. and Y. Zenou (2016), "Racial identity and education in social networks," *Social Networks* 44, 85-94.

[41] Rodriguez, R. (1982), *Hunger of Memory: The Education of Richard Rodriguez*, New York: Bantam.

[42] Schelling, T.C. (1971), "Dynamic models of segregation," *Journal of Mathematical Sociology* 1, 143-186.

[43] Stack, C.B. (1976), *All Our Kin: Strategies for Survival in a Black Community*, New York: Harper & Row.

[44] Tarbush, B. and A. Teytelboym (2014), "Friending," Unpublished manuscript, MIT.

[45] Zhang, J. (2004), "Residential segregation in an all-integrationist world," *Journal of Economic Behavior and Organization* 54, 533-550.

APPENDIX: PROOFS

Proof of Proposition 1.

Before proving this proposition, let us state a useful lemma, which shows under which network structure the condition $c < \delta - \delta^2$ guarantees that each community will be fully intra-connected.

Lemma 1 *If each green and each blue individual has at most one link with the other community, then $c < \delta - \delta^2$ guarantees that all equilibrium networks are such that each community is fully intra-connected.*

Proof: The proof of this lemma is straightforward. Indeed, when each individual has at most one inter-community link, her rate of exposure does not change when adding an intra-community link since it is always equal 1. As a result, the benefit of creating an intra-community link is always $\delta - \delta^2 - c$, which is positive if $c < \delta - \delta^2$. Thus, it is always beneficial to be linked to all individuals from the same community. ∎

Let us now prove Proposition 1.

(i) *Complete integration* between communities: We want to check that this is a pairwise stable equilibrium. We cannot use Lemma 1 since each individual has more than one link with the other community. We need, however, to assume that $c < \delta - \delta^2$, otherwise fully intra-connected communities cannot be studied. Indeed, $c < \delta - \delta^2$ is a necessary condition for fully intra-connected communities but not a necessary and sufficient condition (see the example of the network displayed in Figure 4b). Observe that, in the complete integration case, since all individuals are connected, they cannot form new links. So we need to check only that nobody wants to sever a link.

The green community:

• There is *no* gain in utility for a green person to sever a link with a green person if:

$$\delta^2 - \delta + c + n^B \left[\left(\frac{n^G - 1}{n - 2} \right) \left(\frac{n^B - 1}{n - 2} \right) - \left(\frac{n^G - 2}{n - 3} \right) \left(\frac{n^B - 1}{n - 2} \right) \right] C < 0$$

Indeed, the first term $\delta^2 - \delta + c$ is the net benefit from severing a link to a green individual because of not having a direct connection and instead having an indirect one with this green person and not paying the intra-community cost c. By severing this link, the green individual also *reduces* her exposure rate from $\frac{n^G - 1}{n-2}$ to $\frac{n^G - 2}{n-3}$ but the rate of exposure of all the blue persons she is linked to is not affected. As a result, the gain of this person in terms of inter-community costs is given by the last term of the left-hand side of the above inequality. Rearranging the condition above leads to (using the fact that $n^G = n - n^B$):

$$\delta - \delta^2 - c > \frac{n^B \left(n^B - 1 \right)^2}{\left(n - 2 \right)^2 \left(n - 3 \right)} C \tag{23}$$

• There is *no* gain in utility for a green person to sever a link with a blue person if:

$$\delta^2 - \delta + c + \left(\frac{n^G - 1}{n - 2}\right)\left(\frac{n^B - 1}{n - 2}\right)C - \left(n^B - 1\right)\left(\frac{n^B - 1}{n - 2}\right)\left[\left(\frac{n^G - 1}{n - 3}\right) - \left(\frac{n^G - 1}{n - 2}\right)\right]C < 0$$

As before, by deleting her link, her benefit decreases and is equal to $\delta^2 - \delta < 0$. To understand the rest of this expression, observe that the rate of exposure of this green person now *increases* from $\frac{n^G-1}{n-2}$ to $\frac{n^G-1}{n-3}$ while the rate of exposure of all the blue persons stay the same and equal to $\frac{n^B-1}{n-2}$. As a result, the direct gain in deleting a link to a blue person is $c + \left(\frac{n^G-1}{n-2}\right)\left(\frac{n^B-1}{n-2}\right)C$ but the loss in terms of increase of linking costs to all the other blue agents is $\left(n^B - 1\right)\left(\frac{n^B-1}{n-2}\right)\left[\left(\frac{n^G-1}{n-3}\right) - \left(\frac{n^G-1}{n-2}\right)\right]C$. Rearranging the inequality above leads to:

$$\delta - \delta^2 - c > \frac{\left(n^B - 1\right)\left(n^G - 1\right)\left(n^G - 2\right)}{(n - 3)(n - 2)^2}C \tag{24}$$

The blue community:

The reasoning is similar; one has to replace "green" by "blue". We have:

• There is *no* gain in utility for a blue person to sever a link with a blue person if:

$$\delta - \delta^2 - c > \frac{n^G\left(n^G - 1\right)^2}{(n - 2)^2(n - 3)}C \tag{25}$$

• There is *no* gain in utility for a blue person to sever a link with a green person if:

$$\delta - \delta^2 - c > \frac{\left(n^G - 1\right)\left(n^B - 1\right)\left(n^B - 2\right)}{(n - 3)(n - 2)^2}C \tag{26}$$

We need that the four conditions (23), (24), (25) and (26) to be satisfied in order to have a complete integration between communities. Remember that we have assumed that $n^G \geq n^B$. As a result, if condition (25) is satisfied, then condition (23) is automatically satisfied. Similarly, if condition (24) is satisfied, then condition (26) is automatically satisfied. We are left with two conditions, (25) and (24), which we rewrite for clarity:

$$\delta - \delta^2 - c > \frac{n^G\left(n^G - 1\right)^2}{(n - 2)^2(n - 3)}C$$

$$\delta - \delta^2 - c > \frac{\left(n^B - 1\right)\left(n^G - 1\right)\left(n^G - 2\right)}{(n - 3)(n - 2)^2}C$$

Since

$$\frac{n^G\left(n^G - 1\right)^2}{(n - 2)^2(n - 3)}C > \frac{\left(n^B - 1\right)\left(n^G - 1\right)\left(n^G - 2\right)}{(n - 3)(n - 2)^2}C$$

the only condition left is (25), which is condition (6) in the Proposition.

(*ii*) Let us show that *complete segregation* between communities is an equilibrium network. Individuals can now delete or create a link.

Because of Lemma 1, we know that the condition $\delta - \delta^2 - c \geq 0$ guarantees that each community is fully intra-connected. So we only need to check inter-community deviations.

• There is *no* gain in utility for a green person to establish a link with a blue person, who is necessarily connected to the rest of the blue community, if:

$$\delta + \left(n^B - 1\right)\delta^2 - (c + C) < 0$$

• Similarly, there is *no* gain in utility for a blue person to establish a link with a green individual, who is necessarily connected to the rest of the green community, if:

$$\delta + \left(n^G - 1\right)\delta^2 - (c + C) < 0$$

Since $n^G \geq n^B$, and because mutual consent is necessary, then condition (7) in the Proposition guarantees that there is complete segregation. ∎

Proof of Proposition 2

(*i*) Let us show that the network described in Figure 5 (where circle and square nodes correspond to green and blue individuals, respectively) is an equilibrium network. Because of Lemma 1, since no individual has more than one inter-community link, we know that the condition $\delta - \delta^2 - c > 0$ guarantees that each community is fully intra-connected. So we only need to check inter-community deviations.

• There is *no* gain in utility for the green person G_1 (see Figure 5) to sever a link with the blue person B_1 if:

$$-\delta - \left(n^B - 1\right)\delta^2 + c + C < 0 \tag{27}$$

• There is *no* gain in utility for the blue person B_1 (see Figure 5) to sever a link with the green person G_1 if:

$$-\delta - \left(n^G - 1\right)\delta^2 + c + C < 0 \tag{28}$$

Since $n^G \geq n^B$, the first condition is more restrictive than the second. Mutual consent in link formation imposes that both conditions have to be satisfied at the same time, hence (27) is a requirement for the network to be pairwise stable.

We have now to check that G_1 and B_1 have no incentives to form a link with other agents than B_1 and G_1, respectively.

• There is *no* gain in utility for the green person G_1 to form a link with a blue agent, who is not B_1, if

$$\delta - \delta^2 - c + \left[1 - \frac{\left(n^G - 1\right)}{n^G}\right] C - \frac{\left(n^G - 1\right)}{n^G} C < 0$$

which is equivalent to:

$$\delta - \delta^2 - c < \left(\frac{n^G - 2}{n^G}\right) C \tag{29}$$

• By symmetry, there is *no* gain in utility for the blue agent B_1 to form a link with a green agent, who is not G_1, if:

$$\delta - \delta^2 - c < \left(\frac{n^B - 2}{n^B}\right) C \tag{30}$$

Because of mutual consent and since $n^G \geq n^B$ only condition (30) is required.

Let us now analyze the green and blue agents other than G_1 and B_1.

• There is *no* gain in utility for any green agent different than G_1 to form a link with B_1 if:

$$\delta + \left(n^B - 1\right)\delta^2 - \left[\delta^2 + \left(n^B - 1\right)\delta^3\right] - \left[c + \left(\frac{n^B - 1}{n^B}\right) C\right] < 0$$

which is equivalent to

$$\frac{n^B}{(n^B - 1)}\left[\delta + \left(n^B - 2\right)\delta^2 - \left(n^B - 1\right)\delta^3 - c\right] < C \tag{31}$$

• By symmetry, there is *no* gain in utility for any blue agent different than B_1 to form a link with G_1 if:

$$\frac{n^G}{(n^G - 1)}\left[\delta + \left(n^G - 2\right)\delta^2 - \left(n^G - 1\right)\delta^3 - c\right] < C \tag{32}$$

Because of mutual consent and since $n^G \geq n^B$, $\delta - \delta^2 - c > 0$ and $\delta < 1$, only condition (31) is required.

• There is *no* gain in utility for any green agent different than G_1 to form a link with a blue other than B_1 if:

$$\delta + \left(n^B - 2\right)\delta^2 - \left(n^B - 1\right)\delta^3 - (c + C) < 0 \tag{33}$$

• By symmetry, there is *no* gain in utility for any blue agent different than B_1 to form a link with a green agent other than G_1 if:

$$\delta + \left(n^G - 2\right)\delta^2 - \left(n^G - 1\right)\delta^3 - (c + C) < 0 \tag{34}$$

Because of mutual consent and since $n^G \geq n^B$, only condition (34) is required.

Let us gather all the five conditions together, which are $\delta - \delta^2 - c > 0$, (27), (30), (31) and (34). For the sake of the exposition, let us write them down. We have:

$$\delta - \delta^2 - c > 0$$

$$C < \delta - \delta^2 - c + n^B\delta^2$$

$$\frac{n^B}{(n^B - 2)}\left(\delta - \delta^2 - c\right) < C$$

$$\frac{n^B}{(n^B-1)}\left[\delta-\delta^2-c+\left(n^B-1\right)\left(\delta^2-\delta^3\right)\right]<C$$
$$\delta-\delta^2-c+\left(n^G-1\right)\left(\delta^2-\delta^3\right)<C$$

Denote $\Delta \equiv \delta-\delta^2-c$. Then conditions (5) and (8) in the Proposition summarize these five inequalities.

(*ii*) Let us show that the network described in Figure 6 is an equilibrium network. Because of Lemma 1, since no individual has more than one inter-community link, we know that the condition $\delta-\delta^2-c>0$ guarantees that each community is fully intra-connected. So we only need to check inter-community deviations. Observe that, because $n^G \geq n^B$, each blue agent has one (and only one) bridge link with a green agent while only each of n^B green agents has one (and only one) link with a blue agent. As a result, we can differentiate between green individuals with a bridge link and without a bridge link while this is not the case for blue agents since they all have a bridge link.

Let us start with *link deletion*:

• There is *no* gain in utility for a green agent with a bridge link to sever this link if:

$$-\delta+\delta^3+c+C<0 \tag{35}$$

• Because of symmetry, this same condition ensures that a blue agent with a bridge link does not have incentives to sever it.

Let us now analyze *link creation*:

• There is *no* gain in utility for a green agent with a bridge link to form a link with a blue agent if:

$$\delta-\delta^2-c-\left(\frac{n^G-1}{n^G}\right)\left(\frac{n^B-1}{n^B}\right)C+\left[1-\left(\frac{n^G-1}{n^G}\right)\right]C<0$$

which is equivalent to

$$\frac{n^G n^B\left(\delta-\delta^2-c\right)}{\left(n^G-1\right)\left(n^B-1\right)-n^B}<C \tag{36}$$

• Because of symmetry, a blue agent does not have incentives to build a new bridge link with a green that has already a bridge link if:

$$\frac{n^G n^B\left(\delta-\delta^2-c\right)}{\left(n^G-1\right)\left(n^B-1\right)-n^G}<C \tag{37}$$

Because of mutual consent and since $n^G \geq n^B$, only (37) is needed.

• A blue agent with a bridge link does not have incentives to build a link with a green that does not have a bridge link if:

$$\delta-\delta^2-c-\left(\frac{n^B-1}{n^B}\right)C+\left(1-\frac{n^B-1}{n^B}\right)C<0$$

which is equivalent to

$$\left(\frac{n^B}{n^B-2}\right)\left(\delta-\delta^2-c\right)<C \tag{38}$$

• A green that has no bridge link does not have incentives to create a link with a blue agent if:

$$\left(\frac{n^B}{n^B-1}\right)\left(\delta-\delta^2-c\right)<C \tag{39}$$

Because of mutual consent and since $n^G \geq n^B$, only (38) is needed.

Let us gather all the four conditions together, which are $\delta-\delta^2-c>0$, (35), (37) and (38). For the sake of the exposition, let us write them down. We have:

$$\delta-\delta^2-c>0$$

$$-\delta+\delta^3+c+C<0$$
$$\frac{n^G n^B\left(\delta-\delta^2-c\right)}{\left(n^G-1\right)\left(n^B-1\right)-n^G}<C$$
$$\left(\frac{n^B}{n^B-2}\right)\left(\delta-\delta^2-c\right)<C$$

It is easily verified that

$$\frac{n^G n^B\left(\delta-\delta^2-c\right)}{\left(n^G-1\right)\left(n^B-1\right)-n^G}>\left(\frac{n^B}{n^B-2}\right)\left(\delta-\delta^2-c\right)$$

and thus we end up with the following conditions:

$$\delta-\delta^2-c>0$$
$$\frac{n^G n^B\left(\delta-\delta^2-c\right)}{\left(n^G-1\right)\left(n^B-1\right)-n^G}<C<\delta-\delta^3-c$$

which corresponds to (5) and (9) where $\Delta \equiv \delta-\delta^2-c$.

(*iii*) Let us show that the network described in Figure 7 is an equilibrium network. We denote B_m, the blue agent that has a bridge link with each of the members of the green community (see Figure 7). We cannot use Lemma 1, since B_m has more than one inter-community link.

Let us start with *link deletion*:

• The blue individual B_m does not want to sever any of her bridge links if:

$$-\delta+\delta^2+c+\left(\frac{n^B-1}{n-2}\right)C-\left(n^G-1\right)\left[\left(\frac{n^B-1}{n-3}\right)-\left(\frac{n^B-1}{n-2}\right)\right]C<0$$

which is equivalent to

$$C<\frac{(n-2)(n-3)}{\left(n^B-1\right)\left(n^B-2\right)}\left(\delta-\delta^2-c\right) \tag{40}$$

• The blue individual B_m does not want to sever any of her intra-community links (i.e. with another blue agent) if:

$$-\delta+\delta^2+c+\left(n^G-1\right)\left[\left(\frac{n^B-1}{n-2}\right)-\left(\frac{n^B-2}{n-3}\right)\right]C<0$$

which is equivalent to:

$$C<\frac{(n-2)(n-3)}{\left(n^G-1\right)^2}\left(\delta-\delta^2-c\right) \tag{41}$$

It is easily verified that

$$\frac{(n-2)(n-3)}{\left(n^G-1\right)^2}<\frac{(n-2)(n-3)}{\left(n^B-1\right)\left(n^B-2\right)}$$

If condition (41) is satisfied, then (40) is automatically satisfied. Thus, only condition (41) is required.

• A green agent does not want to sever her bridge link with the blue B_m if:

$$-\delta+\delta^2-\left(n^B-1\right)\left(\delta^2-\delta^3\right)+c+\left(\frac{n^B-1}{n-2}\right)C<0$$

which is equivalent to

$$C<\left(\frac{n-2}{n^B-1}\right)\left[\delta-\delta^2-c+\left(n^B-1\right)\left(\delta^2-\delta^3\right)\right] \tag{42}$$

• It is easily verified that the condition $\delta-\delta^2-c>0$ guarantees that a green agent will never delete a link with another green agent and that a blue agent B_0, which is not B_m, will never delete a link with another blue agent B_0.

Let us now analyze *link creation*:

• Any of the agents denoted by B_0 does not have incentives to directly connect with a green agent if:

$$\delta-\delta^2-c-\left(\frac{n^G-1}{n^G}\right)C<0$$

which is equivalent to

$$C>\left(\frac{n^G}{n^G-1}\right)\left(\delta-\delta^2-c\right) \tag{43}$$

• A green agent does not have incentives to connect with a blue B_0 if:

$$\delta-\delta^2-c-\left(\frac{n^G-1}{n^G}\right)C+\left[\left(\frac{n^B-1}{n-2}\right)-\left(\frac{n^G-1}{n^G}\right)\left(\frac{n^B-1}{n-2}\right)\right]C<0$$

which is equivalent to

$$C>\frac{n^G(n-2)}{\left[\left(n^G-1\right)(n-2)-\left(n^B-1\right)\right]}\left(\delta-\delta^2-c\right) \tag{44}$$

Let us gather all the five conditions together, which are $\delta - \delta^2 - c > 0$, (41), (42), (43) and (44). For the sake of the exposition, let us write them down. We have:

$$\delta - \delta^2 - c > 0$$
$$C < \frac{(n-2)(n-3)}{(n^G-1)^2}\left(\delta - \delta^2 - c\right)$$
$$C < \left(\frac{n-2}{n^B-1}\right)\left[\delta - \delta^2 - c + \left(n^B - 1\right)\left(\delta^2 - \delta^3\right)\right]$$
$$C > \left(\frac{n^G}{n^G-1}\right)\left(\delta - \delta^2 - c\right)$$
$$C > \frac{n^G(n-2)}{\left[(n^G-1)(n-2) - (n^B-1)\right]}\left(\delta - \delta^2 - c\right)$$

It is easily verified that

$$\frac{n^G}{n^G-1} < \frac{n^G(n-2)}{(n^G-1)(n-2) - (n^B-1)}$$

Thus these five conditions reduce to (5) and (10) in the proposition. ∎

Proof of Proposition 3

Consider the bipartite network described in Figure 9 where circle and square nodes correspond to green and blue individuals, respectively. There are n^G individuals, each being connected to all the n^B blue agents. There are n^B individuals, each being connected to all the n^G green agents.

Let us start with *link deletion*:

• A green agent does not have incentives to sever a link with a blue agent if:

$$-\delta + \delta^3 + c < 0 \tag{45}$$

Observe that, here, the inter-community cost C does not appear in (45) because the exposure rate of each green agent is zero and thus the inter-community cost is equal to the intra-community cost, which is c.

• A blue agent does not have incentives to sever a link with a green agent if:

$$-\delta + \delta^3 + c < 0$$

which is (45).

Let us now study *link creation*:

• A green agent does not have incentives to create a link with another green agent if:

$$\delta - \delta^2 - c < 0 \tag{46}$$

Observe that when the green agent creates a link with another green agent, her exposure increases from 0 to $1/n^B$ but the exposure rate of all the blue agents is still 0 (they have no link with each other) and thus the inter-community cost of the green agent is still c and does not increase.

• A blue agent does not have incentives to create a link with another blue agent if:

$$\delta - \delta^2 - c < 0$$

which is (46).

To summarize, we have two conditions that guarantee that the bipartite network displayed in Figure 9 is an equilibrium network, which are (45) and (46). Putting them together, we obtain (11). ∎

Proof of Proposition 4

(*ii*) Let us first show that *complete integration* between communities is always an equilibrium network. In fact, we can use the proof of Proposition 1 (*i*) and add a k when necessary to end up with the four following conditions

$$\delta - \delta^2 - c > \left[\frac{n^B\left(n^B-1\right)^2}{(n-2)^2(n-3)}\right]C$$

$$\delta - \delta^2 - c > \left[k + \frac{\left(n^B-1\right)\left(n^G-1\right)\left(n^G-2\right)}{(n-3)(n-2)^2}\right]C$$

$$\delta - \delta^2 - c > \left[\frac{n^G\left(n^G-1\right)^2}{(n-2)^2(n-3)}\right]C$$

$$\delta - \delta^2 - c > \left[k + \frac{\left(n^G-1\right)\left(n^B-1\right)\left(n^B-2\right)}{(n-3)(n-2)^2}\right]C$$

Because $n^G \geq n^B$, we end up with two conditions, which are

$$\delta - \delta^2 - c > \left[\frac{n^G\left(n^G-1\right)^2}{(n-2)^2(n-3)}\right]C$$

$$\delta - \delta^2 - c > \left[k + \frac{\left(n^B-1\right)\left(n^G-1\right)\left(n^G-2\right)}{(n-3)(n-2)^2}\right]C$$

which are equivalent to (14) in part (*ii*) of the proposition.

(*iii*) Let us now show that *complete segregation* between communities is an equilibrium network. Again, using the proof of Proposition 1 (*ii*) and adding the k when necessary, the condition to have complete segregation is given by:

$$C > \frac{\delta + \left(n^G-1\right)\delta^2 - c}{1+k}$$

which is (15) in part (*iii*) of the proposition.

(*i*) Let us find the condition that guarantees that there are no equilibrium for which each community is not fully connected (i.e. bipartite networks). For that, we take the worst case scenario. The smallest benefit a blue person can obtain by making a link to another blue is $\delta - \delta^2$. The highest cost for a blue i to have a link with another blue is equal to:

$$\min_b \left\{ -c + n^G \left[\frac{b}{n^G + b} \times 1 - \frac{b+1}{n^G + b + 1} \times 1 \right] C \right\}$$

where $b \in \left[0, n^B - 2\right]$ is the number of blue friends of blue i. Observe that

$$\frac{b}{n^G + b} - \frac{b+1}{n^G + b + 1} = -\frac{n^G}{(n^G + b + 1)(n^G + b)} < 0$$

So

$$\min_b \left[\frac{b}{n^G + b} - \frac{b+1}{n^G + b + 1} \right] \Leftrightarrow b = 0$$

This implies that the worst case scenario is

$$\delta - \delta^2 - c - \frac{n^G}{(n^G + 1)} C > 0$$

$$\Leftrightarrow C < \frac{\left(n^G + 1\right)\left(\delta - \delta^2 - c\right)}{n^G}$$

If this is true then any blue will create a link with another blue. Doing the same procedure for greens, we obtain

$$C < \frac{\left(n^B + 1\right)\left[\delta - \delta^2 - c\right]}{n^B}$$

Since

$$\frac{\left(n^G + 1\right)\left[\delta - \delta^2 - c\right]}{n^G} < \frac{\left(n^B + 1\right)\left[\delta - \delta^2 - c\right]}{n^B}$$

Then the condition for both blues and greens is

$$C < \frac{\left(n^G + 1\right)\left(\delta - \delta^2 - c\right)}{n^G}$$

which is (13) in part (*i*) of the proposition. ∎

Proof of Proposition 5

In Figure 10 (*i*, *ii*, *iii* and *iv*), circle and square nodes correspond to green and blue individuals, respectively.

(*i*) Let us show that the network described in Figure 10 (*i*) is an equilibrium network.

Let us start with *link deletion*:

• Any green agent does not have incentives to sever a link with another green agent if:

$$c < \delta \tag{47}$$

• Any blue agent does not have incentives to sever a link with another blue agent if:

$$c < \delta$$

which is (47).

Let us now consider *link creation*:

• The center in the star formed by the green community (referred to as the "green center") does not have incentives to build a link with the center in the star formed by the blue community (referred to as the "blue center") if:

$$\delta + \left(n^B - 1\right)\delta^2 - c < C \tag{48}$$

• Similarly, the blue center does not have incentives to build a link with the green center if:

$$\delta + \left(n^G - 1\right)\delta^2 - c < C \tag{49}$$

Because of mutual consent and because $n^G \geq n^B$, only (48) is required (since it implies (49)).

• If the centers have no incentive to connect to each other, *a fortiori* no individual from one community has incentives to connect with an individual from the other community.

• Finally, we need to consider the possible link creation between two green individuals (both are not the center) or between two blue individuals (both are not the center). The condition for this not to happen is:

$$\delta - \delta^2 - c < 0 \tag{50}$$

We are thus left with three conditions (47), (48) and (50). Combining (47) and (50) gives (16) while condition (48) is (17) in part (i) of the proposition.

(ii) Let us show that the network described in Figure 10 (ii) is an equilibrium network.

Let us start with *link deletion*:

• The green center agent does not have incentives to sever the bridge link with the blue center if:

$$-\delta - \left(n^B - 1\right)\delta^2 + c - C < 0 \tag{51}$$

• The blue center agent does not have incentives to sever the bridge link with the green center if:

$$-\delta - \left(n^G - 1\right)\delta^2 + c - C < 0 \tag{52}$$

Because of mutual consent and because $n^G \geq n^B$, only (51) is required (since it implies (52)).

• None of the centers has incentives to sever a link with her own community if

$$c < \delta \tag{53}$$

because they have just one bridge link with the other community, and we can once more apply the result that there is no *exposure effect* in this case.

• A green agent does not have incentives to sever the link with the green center if:

$$-\delta - \delta^2 - \left(n^B - 1\right)\delta^3 + c < 0$$

When (53) holds, then this inequality is always satisfied.

• A blue agent does not have incentives to sever the link with the blue center if:

$$-\delta - \delta^2 - \left(n^G - 1\right)\delta^3 + c < 0$$

When (53) holds, then this inequality is always satisfied.

Let us now study *link creation*:

• The green center does not have incentives to connect with a peripheral agent of the blue community if:

$$\delta - \delta^2 - c - \left(\frac{n^G - 1}{n^G}\right)C + \left[1 - \left(\frac{n^G - 1}{n^G}\right)\right]C < 0$$

which is equivalent to:

$$C > \frac{n^G}{(n^G - 2)}\left(\delta - \delta^2 - c\right) \tag{54}$$

• Similarly, the blue center does not have incentives to connect with a peripheral agent of the green community if:

$$C > \frac{n^B}{(n^B - 2)}\left(\delta - \delta^2 - c\right) \tag{55}$$

Because of mutual consent and because $n^G \geq n^B$, only (55) is required (since it implies (54)).

• Since (55) ensures that none of the centers have incentives to form a link with the periphery of the other community, and since mutual consent is necessary for link formation, we don't have to check for the conditions that ensure that a peripheral agent does not have incentives to connect with the center of the other community.

• A green peripheral agent does not have incentives to connect to a blue peripheral if:

$$\delta - \delta^3 - c < C \tag{56}$$

• Because of symmetry, condition (56) ensures that a blue peripheral agent does not have incentives to connect with a green peripheral agent.

• Finally, we need to consider the possible link creation between two green peripheral individuals or between two blue peripheral individuals. The condition for this not to happen is:

$$\delta - \delta^2 - c < 0 \tag{57}$$

To summarize, we are left with five conditions (51), (53), (55), (56) and (57). These conditions are:

$$C > c - \delta - \left(n^B - 1\right)\delta^2$$

$$c < \delta$$

$$C > \frac{n^B}{(n^B - 2)}\left(\delta - \delta^2 - c\right)$$

$$C > \delta - \delta^3 - c$$

$$c > \delta - \delta^2$$

It is easily verified that

$$\delta - \delta^3 - c > \frac{n^B}{(n^B - 2)}\left(\delta - \delta^2 - c\right)$$

and

$$\delta - \delta^3 - c > c - \delta - \left(n^B - 1\right)\delta^2$$

As a result, we are left with $\delta - \delta^2 < c < \delta$ and $C > \delta - \delta^3 - c$, which correspond to (16) and (18) in part (*ii*) of the proposition.

(*iii*) Let us show that the network described in Figure 10 (*iii*) is an equilibrium network.

Let us start with *link deletion*:

• A green peripheral agent (with a bridge) does not have incentives to sever her bridge link with a blue peripheral if:

$$-\delta + \delta^5 - \delta^2 + \delta^4 - c - C < 0 \tag{58}$$

• Condition (58) also ensures that a blue peripheral agent does not have incentives to sever her bridge link with a green peripheral agent.

• None of the centers has incentives to sever a link with her own community if

$$-\delta + c - \delta^2 + \delta^4 < 0$$

The condition

$$c < \delta \tag{59}$$

guarantees that the inequality above is always true.

Let us now study *link creation*:

Observe that because $n^G \geq n^B$, then all blue agents but the center have one bridge link while only n^B green agents (which does not include the green center) have one bridge link.

• A green peripheral agent without a bridge does not have incentives to form a link with another green peripheral agent without a bridge if:

$$\delta - \delta^2 < c \tag{60}$$

• A green peripheral agent with a bridge does not have incentives to form a link with another green peripheral agent without a bridge if:

$$\delta - \delta^2 < c$$

which is (60).

• Because of mutual consent, if (60) holds, then a green peripheral agent without a bridge cannot form a link with another green peripheral agent with a bridge.

• A green peripheral agent with a bridge does not have incentives to form a link with another green peripheral agent with a bridge if:

$$\delta - \delta^2 + \delta^2 - \delta^3 < c$$

which is equivalent to

$$\delta - \delta^3 < c \tag{61}$$

• Because of symmetry, condition (61) also ensures that a blue peripheral agent does not have incentives to form a link with another blue peripheral agent.

• The green center does not have incentives to form a link with the blue center if:

$$\delta - \delta^3 - c < C \tag{62}$$

• Because of symmetry, condition (62) also ensures that the blue center does not have incentives to form a link with the green center.

Observe that if (61) holds, then (62) immediately follows. As a result, only (60) is required.

• The green center does not have incentives to form a link with a blue peripheral agent if:

$$\delta - \delta^2 + \delta^2 - \delta^3 - c < \frac{C}{2}$$

which is equivalent to:

$$\delta - \delta^3 - c < \frac{C}{2}$$

which holds as long as that (61) holds.

• Because of symmetry, the same condition ensures that the blue center does not have incentives to form a link with a green peripheral agent with a bridge.

• The blue center does not have incentives to form a link with a green peripheral agent without a bridge if:

$$\delta - \delta^4 + \delta^2 - \delta^3 + \delta^3 - \delta^4 - c < C$$

which is equivalent to:

$$\delta - 2\delta^4 + \delta^2 - c < C \tag{63}$$

• Because of symmetry, this same condition ensures that a green peripheral agent without a bridge does not have incentives to form a link with the blue center.

• The blue center does not have incentives to form a link with a green peripheral agent with a bridge if:

$$\delta - \delta^2 + \delta^2 - \delta^3 + \delta^3 - \delta^4 - c < \frac{C}{2}$$

which is equivalent to:

$$2\left(\delta - \delta^4 - c\right) < C \tag{64}$$

• Because of symmetry, this same condition ensures that a green peripheral agent with a bridge does not have incentives to form a link with the blue center.

• A green peripheral agent with a bridge does not have incentives to form a (second) bridge link with another blue peripheral agent iff

$$\delta - \delta^3 - c - \frac{1}{4}C + \frac{1}{2}C < 0$$

which is equivalent to:

$$C < 4\left(c - \delta + \delta^3\right) \tag{65}$$

• Because of symmetry, this same condition ensures that a blue peripheral agent does not have incentives to form a (second) bridge link with a green peripheral agent with a bridge.

• A green peripheral agent without a bridge does not have incentives to form a bridge link with a blue peripheral agent if:

$$\delta - \delta^3 + \delta^2 - \delta^4 - c - \frac{1}{2}C < 0$$

which is equivalent to:

$$C > 2\left(\delta - \delta^3 + \delta^2 - \delta^4 - c\right) \tag{66}$$

• Because of symmetry, this same condition ensures that a blue peripheral agent does not have incentives to form a (second) bridge link with a green peripheral agent without a bridge.

To summarize, we are left with eight conditions (58), (59), (60), (61), (63), (64), (65) and (66). These conditions are:

$$-\delta + \delta^5 - \delta^2 + \delta^4 - c < C$$

$$\delta - \delta^2 < c < \delta$$

$$\delta - \delta^3 < c$$

$$\delta - 2\delta^4 + \delta^2 - c < C$$

$$2\left(\delta - \delta^4 - c\right) < C$$

$$C < 4\left(c - \delta + \delta^3\right)$$
$$C > 2\left(\delta - \delta^3 + \delta^2 - \delta^4 - c\right)$$

It is easily verified that

$$\delta - 2\delta^4 + \delta^2 - c > -\delta + \delta^5 - \delta^2 + \delta^4 - c$$

and

$$2\left(\delta - \delta^3 + \delta^2 - \delta^4 - c\right) > 2\left(\delta - \delta^4 - c\right)$$

Thus, we are left with

$$\delta - \delta^2 < c < \delta$$
$$\delta - \delta^3 < c$$

$$\max\left\{\delta - 2\delta^4 + \delta^2 - c, 2\left(\delta - \delta^3 + \delta^2 - \delta^4 - c\right)\right\} < C < 4\left(c - \delta + \delta^3\right)$$

For this inequality to make sense, it has to be that $2\left(\delta - \delta^3 + \delta^2 - \delta^4 - c\right) < 4\left(c - \delta + \delta^3\right)$, which is true if $c > \delta - \delta^3 + \delta^2/3$, which includes $c > \delta - \delta^3$. It also has to be that $\delta - 2\delta^4 + \delta^2 - c < 4\left(c - \delta + \delta^3\right)$, which is equivalent to:

$$c > \delta - \frac{2}{5}\delta^4 + \frac{\delta^2}{5} - \frac{4}{5}\delta^3$$

As a result, we are left with the following three conditions

$$\delta - \delta^2 < c < \delta$$
$$c > \delta - \frac{2}{5}\delta^4 + \frac{\delta^2}{5} - \frac{4}{5}\delta^3$$
$$\max\left\{\delta - 2\delta^4 + \delta^2 - c, 2\left(\delta - \delta^3 + \delta^2 - \delta^4 - c\right)\right\} < C < 4\left(c - \delta + \delta^3\right)$$

which are stated in part *(iii)* of the proposition.

(iv) Let us show that the network described in Figure 10 *(iv)* is an equilibrium network.

Let us start with *link deletion*:

• The two centers don't have incentives to sever the bridge link that connects them if:

$$-\delta + \delta^3 + c + C < 0 \tag{67}$$

• A blue peripheral individual does not have incentives to sever the link with a green peripheral one iff

$$-\delta + \delta^3 + c + \frac{1}{n^G - 1}\frac{1}{n^B - 1}C - \left(n^G - 2\right)\left[\frac{1}{n^G - 2}\frac{1}{n^B - 1} - \frac{1}{n^G - 1}\frac{1}{n^B - 1}\right]C < 0$$

which is equivalent to

$$\delta - \delta^3 - c > 0$$

which trivially holds if (67) holds too. The same argument holds to show that a green peripheral individual does not have incentives to sever the link with a blue peripheral one.

• A peripheral blue has no incentives to sever the link with the blue center if

$$-\delta + \delta^3 - c + \left(n^G - 1\right) \frac{1}{(n^G - 1)} C < 0$$

which is equivalent to:

$$-\delta + \delta^3 - c + C < 0$$

This inequality is satisfied as long as (67) holds.

Let us study with *link creation*:

• A peripheral blue individual does not have incentives to form a link with the center of the other community if:

$$\delta - \delta^2 - c - \frac{1}{n^G} \frac{n^G - 1}{n^G} C - \left(n^G - 1\right) \left[\frac{1}{n^G} \frac{1}{n^B - 1} - \frac{1}{n^G - 1} \frac{1}{n^B - 1} \right] C < 0$$

which is equivalent to:

$$\delta - \delta^2 - c < \frac{1}{n^G} \left[\frac{n^G - 1}{n^G} - \frac{1}{n^B - 1} \right] C$$

which is a condition that is trivially satisfied given the assumption that $c > \delta - \delta^2$ and that the right hand side of this last inequality is strictly positive.

• An equivalent argument is valid for the incentives of a peripheral green not willing to form a link with the blue center. Hence, because of mutual consent, we do not have to check for the condition of a center of one of the communities not willing to form a link with a peripheral agent of the other community.

• A peripheral agent does not have incentives to build a link with another peripheral of her own community because the direct benefit of this connection would be

$$\delta - \delta^2 - c < 0$$

and it would imply higher costs for the connections with the other community.

Hence, when (67) holds this network is stable. ▮

Proof of Proposition 6

The total surplus for complete segregation is equal to:

$$\left[n^G \left(n^G - 1\right) + n^B \left(n^B - 1\right)\right] (\delta - c)$$

while the total surplus for complete integration is given by:

$$n^G \left[\left(n^G - 1\right)(\delta - c) + n^B \delta - \left(c + \frac{\left(n^G - 1\right)\left(n^B - 1\right)}{(n-1)^2} C \right) \left(n^B - 1\right) \right]$$
$$+ n^B \left[\left(n^B - 1\right)(\delta - c) + n^G \delta - \left(c + \frac{\left(n^G - 1\right)\left(n^B - 1\right)}{(n-1)^2} C \right) \left(n^G - 1\right) \right]$$

Segregation is better if and only if:

$$\delta \leq \left[c + \frac{\left(n^G - 1\right)\left(n^B - 1\right)}{(n-1)^2} C \right] \left(1 - \frac{n}{2n^G n^B}\right) \tag{68}$$

It is easy to check that $1 > \frac{n}{2n^G n^B}$ and, hence, that the upper bound is strictly positive.

For C large enough, segregation dominates integration. What happens when C is smaller? Let's take the smallest value C can take, i.e. $\delta + \left(n^B - 1\right)\delta^2 - c$, and see if integration dominates segregation. The condition (68) is now given by:

$$\delta > \left[c + \frac{\left(n^G - 1\right)\left(n^B - 1\right)\left[\delta + \left(n^B - 1\right)\delta^2 - c\right]}{(n-1)^2} \right] \left(1 - \frac{n}{2n^G n^B}\right)$$

where C has been replaced by $\delta + \left(n^B - 1\right)\delta^2 - c$. This is equivalent to:

$$\delta \left[\frac{2(n-1)^2 n^G n^B + \left(n^G - 1\right)\left(n^B - 1\right)\left(2n^G n^B - n\right)}{2n^G n^B - n} \right]$$
$$> \left[(n-1)^2 - \left(n^G - 1\right)\left(n^B - 1\right) \right] c + \left(n^G - 1\right)\left(n^B - 1\right)\left(n^B - 1\right)\delta^2$$
$$\Leftrightarrow \delta \left[\frac{2n^G n^B}{2n^G n^B - n} - \frac{\left(n^G - 1\right)\left(n^B - 1\right)}{(n-1)^2} \right] - \frac{\left(n^G - 1\right)\left(n^B - 1\right)^2}{(n-1)^2} \delta^2$$
$$> \left[1 - \frac{\left(n^G - 1\right)\left(n^B - 1\right)}{(n-1)^2} \right] c$$
$$\Leftrightarrow c < \delta \left[\frac{\frac{2n^G n^B}{2n^G n^B - n} - \frac{\left(n^G-1\right)\left(n^B-1\right)}{(n-1)^2}}{1 - \frac{\left(n^G-1\right)\left(n^B-1\right)}{(n-1)^2}} \right] - \left[\frac{\frac{\left(n^G-1\right)\left(n^B-1\right)^2}{(n-1)^2}}{1 - \frac{\left(n^G-1\right)\left(n^B-1\right)}{(n-1)^2}} \right] \delta^2$$

We are in the range $c < \delta - \delta^2$. It is easy to verify that:

$$\frac{\frac{2n^G n^B}{2n^G n^B - n} - \frac{\left(n^G-1\right)\left(n^B-1\right)}{(n-1)^2}}{1 - \frac{\left(n^G-1\right)\left(n^B-1\right)}{(n-1)^2}} > 1$$

So, with the help of some algebra, we find that a sufficient condition is that

$$\frac{\frac{\left(n^G-1\right)\left(n^B-1\right)^2}{(n-1)^2}}{1 - \frac{\left(n^G-1\right)\left(n^B-1\right)}{(n-1)^2}} \leq 1$$

which is equivalent to (21). ∎