Submission Number: PET11-11-00137

Identity and Social Distance in Friendship Formation

Joan De marti
Universitat Pompeu Fabra

Yves Zenou
Stockholm University and IFN

Abstract

We analyze a model of network formation with agents that belong to different communities. Both individual benefi ts and costs depend on direct as well as indirect connections. Bene fits of an indirect connection decrease with distance in the network. The cost of a link depends on the type of agents involved in it as well as the the rest of linkage decisions of both of them: two individuals from the same community always face a low linking cost, and the cost of forming a relationship for two individuals belonging to different communities diminishes with the rate of exposure of each of them to the other community. As a result, our model introduces endogenous social distances determined by individuals' identi fication, according to exposure in the resulting network, to each community. We derive a number of results with regard to equilibrium networks. In particular, socialization among the same type of agents might be weak even if the within-type link cost is very low, and oppositional identity patterns can arise for a wide range of parameters. Finally, our model suggests that, if we take into account the cost differences, policies to reduce segregation are going to be effective and socially desirable only if these policies reduce the gap of inter- versus intra-community linking costs by a sufficiently large extent.

We are grateful to the participants of the 13th Coalition Theory Network Workshop, the seminars at IFN, IIES, and Duke University, in particular, Andrea Galeotti, Guillaume Haeringer, Willemien Kets, and Rachel Kranton for helpful comments.
Submitted: February 28, 2011.

Identity and Social Distance in Friendship Formation*

Joan de Marti[†] Yves Zenou[‡]

February 28, 2011

Abstract

We analyze a model of network formation with agents that 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. The cost of a link depends on the type of agents involved in it as well as the the rest of linkage decisions of both of them: two individuals from the same community always face a low linking cost, and the cost of forming a relationship for two individuals belonging to different communities diminishes with the rate of exposure of each of them to the other community. As a result, our model introduces endogenous social distances determined by individuals' identification, according to exposure in the resulting network, to each community. We derive a number of results with regard to equilibrium networks. In particular, socialization among the same type of agents might be weak even if the within-type link cost is very low, and oppositional identity patterns can arise for a wide range of parameters. Finally, our model suggests that, if we take into account the cost differences, policies to reduce segregation are going to be effective and socially desirable only if these policies reduce the gap of inter- versus intra-community linking costs by a sufficiently large extent.

Keywords: networks, identity, homophily, social norms.
JEL Classification: D85, J15.

*We are grateful to the participants of the 13th Coalition Theory Network Workshop, the seminars at IFN, IIES, and Duke University, in particular, Andrea Galeotti, Guillaume Haeringer, Willemien Kets, and Rachel Kranton for helpful comments.

[†]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

The concept of identity has been analyzed for decades in philosophy, psychology, and sociology (see, e.g. Abrams and Hogg, 1999). It is, however, only recently that it has captured the attention of economists. Akerlof and Kranton (2000) were the first to introduce identity into the neoclassical utility maximizing framework in an analysis that draws directly from social psychology's social identity approach and self-categorization theory.[1]

In the present paper, we adopt a different but related view of *identity* by highlighting the importance of *exposure to the other group* in the friendship formation process between individuals of different ethnic groups. We study a network formation game where individuals belong to different communities. The main novelty is that in our model linking decisions determine the endogenous costs and benefits of individual exposure and identification to other communities.

Motivation. Part of the literature has visualized the concept of identity as unidimensional. In other words, individuals with a stronger identification to their own group are usually assumed to have a weaker identification to the other group. Identifications with own and other cultures are treated as mutually exclusive. This has usually been studied in societies where a majority and a minority culture coexist. Those who adopt this view consider that ethnic minorities either remain persistent and loyal to their inherited ethnicity or assimilate to the ethnic environment of the majority group. This can lead to the phenomenon of *oppositional identities*, where some ethnic minorities reject the majority behavioral norms while others totally assimilate to it (Ainsworth-Darnell and Downey, 1998). For example, studies in the US (and also in the UK) have found that African American students in poor areas may be ambivalent about learning standard English and performing well at school because this may be regarded as "acting white" and adopting mainstream identities (Fordham and Ogbu, 1986; Wilson, 1987; Delpit, 1995; Ogbu, 2003; Austen-Smith and Fryer, 2005; Selod and Zenou, 2006; Battu et al., 2007; Fryer and Torelli, 2010; Battu and Zenou, 2010).[2]

There is a literature in psychology (see, in particular, Phinney, 1990; Berry, 1997; Ryder et al., 2000) that proposes a broader concept of self-identification in a *two-dimensional* framework,

[1] For an overview of the literature on the economics of identity, see Kirman and Teschl (2004) and Akerlof and Kranton (2010).

[2] There are few theoretical models that try to explain oppositional identity behaviors. Austen-Smith and Fryer (2005) model these types of trade offs faced by black individuals. They put forward the tension faced by blacks between signalling their type to the outside labor market and signalling their type to their peers: signals that induce high wages can be signals that induce peer rejection. Battu et al. (2007) highlight the trade offs faced by blacks. On the one hand, they want to interact with other blacks and thus to reject the white's norm. On the other, they also want to be friends with whites because the latter possess a higher quality social capital. They find that black workers can end up choosing oppositional identities if their identity is not strong enough or the wage premium of being employed is high enough. Based on cultural transmission and peer effects, Bisin et al. (2010) develop a dynamic model of identity formation that explains why ethnic minorities may choose to adopt oppositional identities and why this behavior may persist over time.

where identifications with two different cultures are not necessary mutually exclusive. Berry (1997) presents four distinct strategies for how individuals relate to two cultures. *Assimilation* is a weak identification with the culture of origin and an strong identification with the alternative culture. *Integration* is achieved when an individual combines strong dedication to the origin and large commitment to the other culture. *Marginalization* is a weak dedication to both cultures. Finally, *separation* is an exclusive commitment to the culture of origin. The following figure summarizes these four different possibilities in a two-dimensional space.

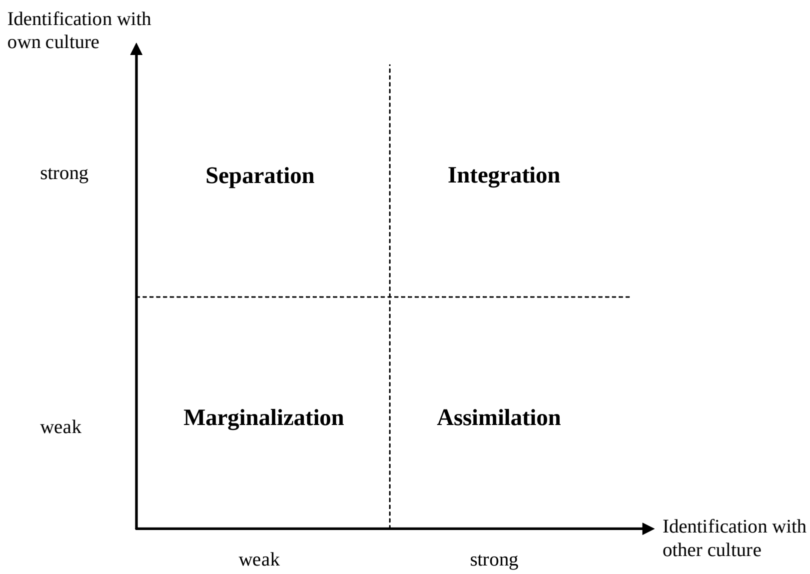
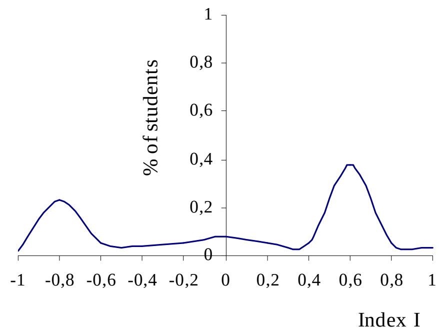


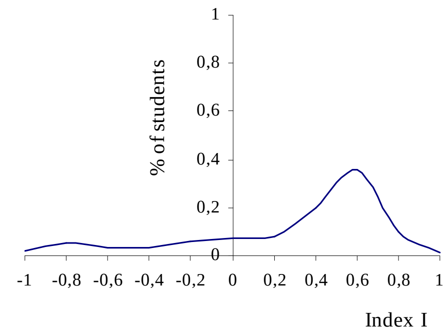
Figure 1. Different identifications for ethnic minorities

As it can be seen from Figure 1, individuals who are integrated have not only a strong identification to the majority culture but also to their own culture. Observe that the previous definition of an oppositional identity corresponds to either a separated or an assimilated individual in Figure 1.

There are some empirical studies in the US using both the unidimensional and bidimensional definition of identity choices. For example, using the National Longitudinal Study of Adolescent Health (AddHealth), Patacchini and Zenou (2007) use the homophily index proposed by Coleman (1958) to analyze the exposure of individuals of white and black race to own and other races. If the homophily index of a student 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 2 displays their results for mixed schools (i.e. schools with a percentage of black and white students between 35 and 75 percent).



Blacks in integrated schools



Whites in integrated schools

Figure 2. Distribution of students by share of same-race friends in integrated schools

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. In terms of Berry's characterization presented above (Figure 1), most white students and some black students show a separated or integrated identities, while a relevant fraction of black students shows assimilated identities.[3]

A model of homogeneous behavior among members of same groups cannot explain the pattern obtained in Figure 2. Choices of friends between races need to be consistent with each other in order for the observed aggregated level of social interactions to show the emergence of heterogeneous identity patterns. Thus, to understand the observed patterns, the *network* aspect of friendships cannot be ignored.

Model and Results. To the best of our knowledge, there are no theoretical models explaining from an strategic point of view both the identity patterns described in Figure 1 and the socialization patterns observed in Figure 2. We propose a network formation model that can simultaneously explain these two aspects. We consider a finite population of individuals composed by two different communities. These two communities are cathegorized according to some exogenous factor such as, for example, sex, race or ethnic and cultural traits. Individuals decide with whom they want to connect according to a utility function that weights the costs and benefits of each connection.

[3] Marmaros and Sacerdote (2006) 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.

The result is a network of relations where there is a link connecting two different individuals only if they are friends. 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.[4] 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 cathegorized according to ethnicity, that 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, mitigate 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.

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 *negatively* depends on the fraction of friends from the other community each person has. This cost is, however, never lower than the cost of intracommunity links.[5]

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, again, 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

[4] See Goyal (2007) and Jackson (2008) for overviews of the growing literature on social and economic networks.

[5] Johnson and Gilles (2000) and Jackson and Rogers (2005) also extend the Jackson and Wolinsky (1996)'s 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.

agents who are not connected have incentives to build a new link. In our model, it is a complex combinatorial problem to fully characterize the set of stable networks, however we provide a partial characterization that conveys information about when different identity and socialization patterns can arise.

In this context, when *intracommunity* linking costs are low, we show that oppositional identities can emerge when *intercommunity* costs are also low, i.e. the maximum possible cost of an intercommunity link is close to the cost of an intracommunity link. 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 positive externalities for the agents who are in their respective neighborhoods since the cost of a link is only paid by the individuals directly involved in it. We can also determine conditions under which totally *assimilated* and *separated* minorities (Figures 1 and 2) can emerge in equilibrium as well as "extreme" networks (i.e. bipartite networks) where individuals of each community are only connected to individuals of the other community.

The mechanism we suggest links socialization costs with network geometry. Since individual and aggregate welfare depend on the geometry of the resulting network, we may wonder about the impact of policies that try to diminish intercommunity socialization costs. Such an analysis is difficult in our context, due to the inherent multiplicity of stable configurations. However, we try to do one step in this direction by comparin two extreme outcomes: extremely integrated and extremely segregated networks. We show that, when intracommunity costs are low, social integration is not always preferred to social segregation. The inefficiency comes from the excessive individual cost payed to build bridge links between communities. This suggests that policies may only be effective if they substantially reduce intercommunity 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) programs implemented in the United States, which relocates families from high- to low-poverty neighborhoods (and from racially segregated to mixed neighborhoods).[6] In our theoretical framework, policies that diminish intercommunity 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 identities, but the outcome is not going to be socially efficient unless these policies sufficiently decrease the cost of interactions.

Our model can be extended in a number of directions. We present two different possible extensions in the last sections of the paper. First, we introduce heterogeneous payoff externalities. It might be that agents of one of the two types exert a larger direct positive externality on others than

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

the other types. This setup can represent, for example, a situation in which one of the two types has *ex ante* a higher human and/or social capital.[7] Second, we introduce social punishment for individuals from the minority group who identifies herself with the majority culture. This punishment expresses the rejection by the members of her original group who strictly stick to their social and cultural values. This can be a reduced form representation of the "acting white" phenomenon mentioned above. Both situations facilitate the adoption of oppositional identities.

Related Literature. The papers by Currarini et al. (2009, 2010), Bramoullé and Rogers (2010), and Mele (2010) study homophily in networks using stochastic models of network formation. The aim in these papers is therefore similar, but there are important differences with respect to the methodology: this set of papers assumes 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 and 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 among the same group of agents. Bramoullé and Rogers (2010) depart from Currarini et al. by assuming that dynamic matching follows the process studied in Jackson and Rogers (2007) and they show that more connected individuals tend to have a more diverse set of friends.[8] Mele (2010) studies a model where meetings are dynamic and stochastic and each individual involved in ameeting can decide whether he wants to create or sever the link with the other person and he shows this process always converges. Mele's model is closer to our one because the utility function he considers also assumes direct and indirect rewards (in his case, only up to distance two). Mele is able to derive precise conclusions about the probability of observing each network, and he uses this to analyze counterfactuals and derive policy implications.

Eguia (2010) presents a theory in which the cost of assimilation is endogenous and strategically chosen by the better-off group in order to screen those who wish to assimilate. Eguia (2010) shows that, in equilibrium, only high types who generate positive externalities to the members of the better-off group will assimilate. The paper does not focus on network issues and therefore the results are of a different and complementary nature.

Some papers analyze the consequences of homophily in social networks. For example, Buhai and Van der Leij (2008) develop a social network model of occupational segregation with inbreeding bias, and Golub and Jackson (2008) 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 patterns. Shelling's model shows that, even mild preferences for interacting with people from the same community can lead to large differences in terms of location decision. Indeed, his results

[7] Benabou (1996) studies a location model with two types and heterogeneous human capital externalities with a similar feature.

[8] They show this holds when students are divided according to sex in high-schools in the AddHealth data set.

suggest that total segregation persists even if most of the population is tolerant about heterogeneous neighborhood composition.[9] Our analysis differs from Schelling's classical framework (and its different extensions) in several directions. First of all, we analyze a network formation game, while in Schelling the network structure is fixed. 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 the 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 natural mechanism of our model (that relates the cost of friendship to the social distance of the 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* (only differing by their attachment to a community) may end up with very different network positions. In particular, *separated*, *integrated*, *marginalized* and/or *assimilated* patterns of friendships (Figure 1) may prevail in equilibrium. Thus, we obtain intragroup asymmetric behaviors in connectivity in a number of equilibrium networks, which allow us to rationalize the friendship patterns observed in Figure 2. With this we don't mean that the result of socialization is always going to lead to segregation and/or oppositional identities, but we show that these patterns can emerge in some circumstances as the result of a *decentralized* process of socialization. . There are 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 collections model in Jackson and Wolinsky (1996), emerges.[10] This is why it is extremely hard, if not impossible, to provide a full-fledged characterization of all possible stable networks.[11]

[9]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 or Zhang, 2004).

[10]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.

[11]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 $N = \{1, \ldots, n\}$. This population is divided into two communities, the *Blue* and the *Green* 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 identity 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: $N_i(g) = \{j \neq i \mid g_{ij} = 1\}$, which is of size $n_i(g)$. The direct contacts of individual i of the same type is $N_i^{\tau(i)}(g) = \{j \neq i, \tau(i) = \tau(j) \mid g_{ij} = 1\}$, and we denote the cardinality of this set by $n_i^{\tau(i)}(g)$.

We present some examples of network configurations. The *circle* is such that each agent has two direct contacts. The *star-shaped network* 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* is such that each agent is in direct relationship with every other agents so that each individual i has $n - 1$ direct contacts.

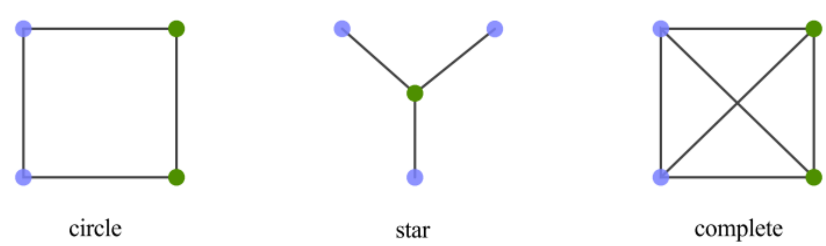


Figure 3. Circle, star and complete networks with four individuals.

A network is depicted as a set of colored nodes (Figure 3), that allow to distinguish among members of different groups, and links that connect some or all of them. Naturally, blue nodes refer to type$-B$ individuals while green nodes indicate type$-G$ individuals.

The circle and the complete network are examples of regular configurations in which all agents share a similar position, though they differ by the number of connections each agents possesses.

The star is an example of centralized, asymmetric, network structure, where the center occupies a very different position than the rest of the other individuals in the network.

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

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.[12] Graphically, there is a path from individual i to individual j whenever one can travel from i to j through the links of the network. The length of a path is the number of links involved in it. The *shortest path* between from 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.

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

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

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 4, 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

[12] 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.

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.

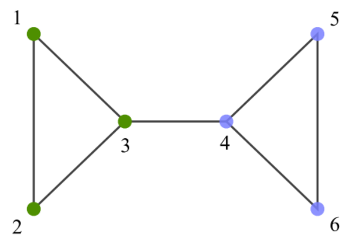


Figure 4. A bridge network.

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 have to introduce first one more concept. Given a network g we define the *rate of exposure* of individual i to her own community as

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

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

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

$$c_{ij}(g) = \begin{cases} c, & \text{if } \tau(i) = \tau(j) \\ c + e_i^{\tau(i)}(g) e_j^{\tau(j)}(g) C, & \text{if } \tau(i) \neq \tau(j) \end{cases} \tag{2}$$

There are thus different costs, depending with whom a connection is made. The main feature in this cost structure is that, since $C > 0$ and rate of exposures are non-negative, it is more costly to form a friendship relationship with someone from the other community (the cost is given by (2)) than with someone from the same community (the cost is c).[13] 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. intercommunity 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

[13] Camargo et al. (2010) show in a randomized experiment that white who are randomly assigned black roommates have in the future a significantly larger proportion of black friends 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 affects significantly economic and social behavior, for example, in forming working relationships.

higher the more different people are. What we have in mind here is that individuals are born with a certain type τ (blue or green) that affects their easiness to 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.[14] But people make choices in terms of friendships, and that be interpreted in terms of identity. These choices can increase or decrease the original gap between individuals. If someone who is born blue chooses to have only blue friends (this is an identity choice) then it will be more difficult for her to interact with a green person. However, the more similar the choices are, the easier is to interact with someone from a different type.[15] Observe that we allow that friend choices can 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 each other than with someone of same type (i.e. the cost is c in both cases).[16]

The reason why we substract a 1 in the denominator in the definition of rate of exposure is because when we compute the cost of a given bridge link between communities we don´t 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 (2), consider again the network described in Figure 4 and assume that individuals 1, 2, and 3 are greens (type G) while individuals 4, 5, and 6 are blues (type B). 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),[17] which implies that $e_3^{\tau(3)}(g) = e_4^{\tau(4)}(g) = 1$.

[14] 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 will create some costs in the interactions between blacks and whites.

[15] 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 find 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 interacting decreases.

[16] In Appendix B, we investigate a different cost function where the intercommunity cost is *not* anymore equal to the intracommunity even if one of the persons involved in a relationship has no friends of the same type.

[17] Observe that, when individual 3 considers the possibility of creating a link with individual 4, individual 3 does

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 be friend to individual 4 (blue) in this situation because the latter has already a green friend.

With the above notation we wanted to highlight that in our model costs, in particular intercommunity costs, depend on the network structure. However, from now on, and to minimize notational burden, we will not make the dependency of the rates of exposure and the linking costs on g explicit.

2.3 Network stability

In games played on 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. Let us now define this concept.

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.[18] It requires that any mutually beneficial link be formed at equilibrium but does not allow for multi-link severance.

We will use throughout this equilibrium concept. 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 assuming that $c < \delta - \delta^2$. If there were only one community (i.e. only one type

not take into account the possible link between 3 and 4 when calculating the percentage of same-race friends of herself and of 4.

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

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 not true anymore: an individual of one type may decide to lower the exposure to his own community to become more attractive to the other one. We start this section by trying to understand under which conditions this may not happen and we still get fully intraconnected communities.

We use the following definitions: A network displays *complete integration* when both communities are completely connected, *complete segregation* when both communities are isolated and *partial integration* in any other case. The first result is the following one. [19]

Proposition 1 *Assume*

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

and that each community is fully intraconnected. Then,

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

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

(*ii*) *If*

$$C > \delta + \left(n^B - 1\right) \delta^2 - c \tag{4}$$

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

It is useful to think about two different effects to interpret the results that we get in the paper. The first effect, the *connections effect*, simply referes to the direct gains and losses of building or severing a link: the gains derive from the benefits that derive from diminishing the distance in the network for the individuals involved in the link, and the cost is simply goven by the cost each of the two sides of the link have to pay to keep it active. This first effect is reminiscent of the use of the connections model payoffs from Jackson and Wolinsky (1996) and, hence, in a sense is not new. However, there is a second effect, the *exposure effect*, that is new. This effect derives from the effect of a new link on the exposure rates of the individuals directly involved in it. If the new link is an intercommunity link, it is going to reduce their respective rates of exposure, and therefore it is going to imply a decrease in the cost of any other intracommunity link in which any of them are directly involved. This indirect effect is in this case positive. If the new link is an intracommunity link, the rates of exposure of the agents involved are going to increase and imply an increase in the cost in their intercommunity links. The indirect exposure effect is in this case negative.

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

Under the light of these two effects, the completely integrated network is going to be stable if the sum of the connections and the exposure effect for any link are positive. Consider an intercommunity link. The connections effect is not clearly signed, because the cost of keeping the link for each of the two sides is strictly large than c, because their rates of exposure to their own communities are strictly positive. However, severing such a link has an strong and negative *exposure effect*: it increases the rate of exposure to their wn communities and the costs for the rest of their intercommunity links increases. Some algebra shows that this second exposure effect always dominates the connections effect and, hence, none of them has incentives to sever the link. The case of an intracommunity link is less clear. In such a case the connections effect is clearly signed: it is positive because we are assuming that $\delta - \delta^2 - c > 0$, which implies that for two individuals from the same community the benefits of a direct connection compared to an indirect connection of distance two outweight the costs of forming such 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 intercommunity links are larger. If C is sufficiently large, then the negative exposure effect overcomes the positive connections effect. That's why we get an upper bound on C in Proposition 1.(i).

The completely segregated network in Proposition 1.(ii) arises when the connections effect of an intercommunity link is negative. The condition in the Proposition is, precisely, the mathematical formulation of this negative condition on the connections effect. Note that in this case, there are no exposure effects to consider since we start from a situation where there are no intercommunity links present. The result sustains in this case on the only existence of connections effects.

Note that, if we use the two-dimensional definition of identity, illustrated in Figure 1, the blues and greens are here *separated*. This could be a case where the two populations are physically separated (i.e. spatially segregated) so that interactions are very costly (because, for example, of commuting costs, prejudices, etc.). Intuitively, if C decreases, individuals may start forming bridge links. These links may make them more attractive, because of the exposure effects, to the other community members, who, in turn, form bridge links, etc.Let us investigate in more details this *partially-integrated* case, where there are some bridges between both communities.

Define

$$\Phi\left(n^{\tau}, \delta, c\right) \equiv \frac{n^{\tau}\left[\delta + \left(n^{\tau} - 2\right)\delta^{2} - \left(n^{\tau} - 1\right)\delta^{3} - c\right]}{n^{\tau} - 1}$$

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 (3) and (**??**).*

(i) If

$$C > \max\left\{\frac{n^{G}\left(\delta - \delta^{2} - c\right)}{n^{G} - 2}, \Phi\left(n^{G}, \delta, c\right), \Phi\left(n^{B}, \delta, c\right)\right\} \tag{5}$$

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

(ii) If

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

holds, then the network where both communities are fully intraconnected 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

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

holds, then the network in which both communities are fully intraconnected and only one blue agent is connected to all the agents of the other community is an equilibrium (Figure 7).

In these equilibrium configurations some integration between blues and greens is taking place. The following figures provide a graphical representation.

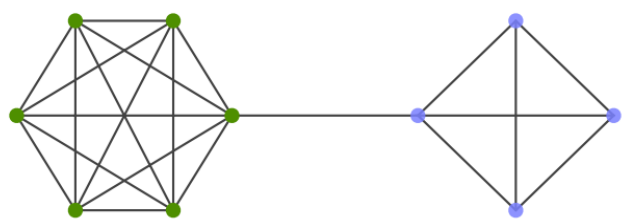


Figure 5. Equilibrium network when condition (5) holds.

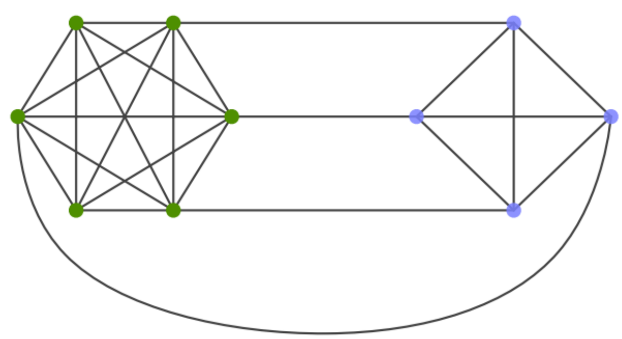


Figure 6. Equilibrium network when condition (6) holds.

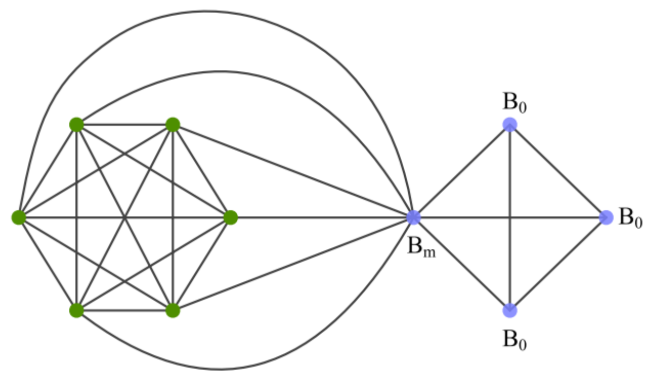


Figure 7. Oppositional identities when $c < \delta - \delta^2$.

To interpret the results, the logic of connections and exposure effects presented above remains. But here a third component becomes more relevant and explicit: mutual consent. Pairwise stability requires that both sides of a link have aligned interests in keeping it active. There is enough that one side prefers to sever it, to actually eliminate the link from the network. When the costs of intercommunity links, parameterized by C, are relatively large, the network in Figure 5 is stable because the connections effect for the agents involved in the only bridge link between communities is positive, but the connections effect of any other intercommunity link is negative for at least one of the two sides of each of these potential links.[20] It is negative because the cost of such connection would be equal to $c + C$, and C is large, and because the already existing bridge links already brings enough externalities from one group to other, and therefore limits the benefits of shortening distances that a new intercommunity link would generate.

When C decreases a bit it is in the interest of individuals from different communities to create one of these missing links, like in Figure 6, because while the direct benefits of a such new connection have not changed, the costs move down, and the sign of the connections effect of such new link reverts. In both the network in Figure 5 and Figure 6, exposure effects play no role, since each of the agents is involved in at most one link, and the cost of this link keeps constant when there are changes in the connections within the community (these intracommunity links do not change the rate of exposure of individuals, that remains maximimal and equal to 1, according to the definition of rate of exposure from the previous section).

The logic behind Figure 7 is different: it strongly relies on the exposure effect. The B_m blue individual invests in a big number of intercommunity links to decrease nough her own rate of exposure, and therefore to decrease her own cost of each of these connections, as well as to make

[20]The two individuals involved in this bridge link enjoy a singular popsition 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 the 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 (Goyal and Vega-Redondo, 2007).

it cheap for each green individual to connect with him and win direct access to the externalities that emerge in the blue community. The positive exposure effect she directly enjoys transforms into a positive connections effect for the other side. This generates the necessary mutual consent to create all these intercommunity links.

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

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

(2) Because links are costly, individuals become more attractive the more they connect to 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. This is highlighted by condition (i) in Definition 1 of the pairwise-stability equilibrium concept. This is *mutual consent*.

Equilibrium networks are those that correctly balance these three forces at the individual level. We hope the equilibrium networks characterized above provide understanding on how these three effects mix with each other. Contrary to the literature on segregation (e.g. Schelling, 1971; Benabou, 1993) and on friendship formation (Austen-Smith and Fryer, 2005; Battu et al., 2007), it is important to observe that both the *individual location* and the *structure of the network* are here crucial to understand the equilibrium outcomes. Indeed, not only benefits but costs are affected by 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.

Let us now investigate the issue of assimilation and oppositional identities. The completely integrated network in Proposition 1 provides a first example of how our model can generate some assimilation patterns: for example, blue individuals partially assimilate with the green community because each blue individual interacts more often with green individuals than with blue individuals. However, in such a case the reason relies mostly on the difference in size of both communities. A more interesting and rich example is the network in Figure 7. Here, accroding to the intuition we gave above, the three effects in our model play a role to generate the assimilation of type B_m. Other stable network could generate similar features, and the reasons behind would be similar: assimilation arises because exposure effect on the side of the B_m blue individual increases the magnitude of the connections effect for each of the green individuals and this induces mutual consent. We believe this interdependency is an important feature to highlight from our model.

With regards to individual welfare, observe that it is not always true that oppositional individuals obtain a higher equilibrium utility than non-oppositional blues. Take, for example, Proposition

?? (Figure 7). The equilibrium utility of the oppositional blue B_m is

$$U_{B_m} = (n-1)(\delta - c) - n^G \left(\frac{n^B - 1}{n-1}\right)\left(\frac{n^G - 1}{n^G}\right) C$$

while the utility of non-oppositional blues is:

$$U_{B_0} = \left(n^B - 1\right)(\delta - c)$$

So we have

$$U_{B_m} \gtreqless U_{B_0}$$
$$\Leftrightarrow C \lesseqgtr \frac{n^G}{n^G - 1} \frac{n-1}{n^B - 1} (\delta - c)$$

This inequality is not incompatible with the condition given in Proposition **??**, meaning that both cases, $U_{B_m} > U_{B_0}$ and $U_{B_m} < U_{B_0}$, are possible. However, if δ is high enough or C or c low enough, then oppositional individuals will be better off. Indeed, on the benefit side, because greens are more numerous, being connected to them give a higher utility to B_m. On the cost side, when C is too high, then B_m is worse off because it is very costly for her to be friend with all the green community. Yet, stability conditions show that even if these links are costly, she is not interested in severing anyone of these bridge links since the benefits she derives from each of them, due to both the connections and the exposure effect, is larger than the cost of keeping one such link active.

3.1.1 Assimilation Patterns and the Exposure Effect

A common feature of all equilibrium networks we have characterized so far is that both communities are fully intraconnected. This limited the type of assimilation and identification patterns we could distinguish there. Now, while we still assume very low intra-community costs, i.e. $c < \delta - \delta^2$, we are going to show that, contrary to the standard connections model, in equilibrium communities can be not fully connected and that oppositional identities and integration can arise. The following proposition characterizes an extreme form of assimilation.

Proposition 3 *Assume (3). If*

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

then the network described in Figure 8, where not fully intraconnected communities prevail and where one blue is assimilated and has an oppositional identity while all other blues are separated, is pairwise stable.

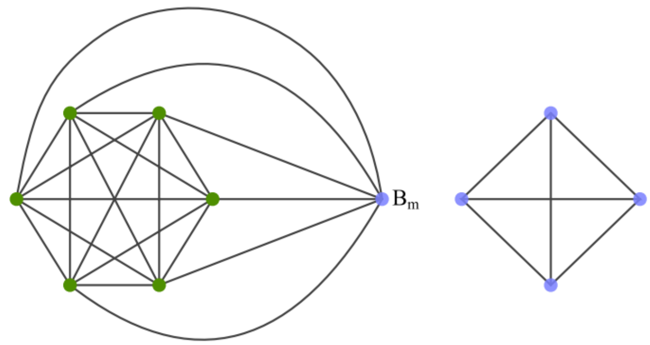


Figure 8. Oppositional identities with non-fully intraconnected communities.

An interesting feature of this equilibrium network is that, while the condition $\delta - \delta^2 - c > 0$ would always induce fully intraconnected communities in a standard connections model, here the blu community is fragmented. the logic behind this result is similar to the one for stability in Proposition 2.(*iii*) (see Figure 7), but taken to the extreme. The B_m blue individual brakes all connections with the blue community to minimize her rate of exposure and eliminate any gap in the cost of all her links with the green community. In a way, this individual fully assimilates and becomes a green individual according to her choice of social connections. This generates an oppositional identity pattern, where a low fraction of blue individuals assimilate with the green community while a majority of blue individuals remain connected with their community of origin. This is much in line with the aggregate conclusions derived from the AddHealth data set we mention in the introduction. Our proposed mechanism provides a rationale for individual social identity choices.

An alternative reading is the following: this result highlights the fact that assimilating to the majority culture (see Figure 1) makes it difficult for a blue person to interact with her own group. In Section 5.2, we further investigate this case by looking at social norms and sanctions where assimilation to the green culture leads to a rejection from the blue community.

Observe that in this network (Figure 8), the blue oppositional B_m has always a higher utility than any other non-oppositional blue B_0 since $(\delta - c)\, n^G > (\delta - c)\, n^B$. Assimilation with the majority brings access to more social externalities.

The previous result shows how it is possible that an agent shows an oppositional identity. The next result shows that, bringing the logic to the extreme, it is even possible that all agents in an economy show an oppositional identity pattern, if C is sufficiently large.

Proposition 4 *Assume (3). If C is sufficiently large, 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 9).*

In the case of a bipartite network each agent is connected only to the other social group and, thus, each agent shows an oppositional identity pattern.

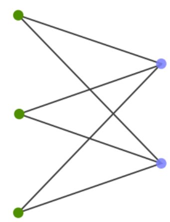


Figure 9. Bipartite Network with $n^W = 3$ and $n^B = 2$.

This network can be sustained in equilibrium because for an individual of a given type, eleiminating all her links with her own community maximizes the positive exposure effect. A link with an agent of same type would be detrimental because while it would be quite inexpensive in direct terms, it would have a negative counterpart: all links with the agents of other type would involve a higher cost, due to the increase in the fraction of same-type friends, or alternatively, due to the decrease in exposure to the other type. This situation can be restated as follows: in a bipartite network, all green agents are "becoming" blues while all blue agents are "becoming" greens.

3.2 Higher socialization costs

Let us now consider the case when $c > \delta - \delta^2$ so that it becomes more expensive to form links with individuals from the same community. 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 individual is always a pairwise stable network.[21] We thus focus on communities that have a star-shaped form. Of course, since we are dealing with a different cost structure, it is 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 form a star network.

Proposition 5 *Assume that*

$$\delta - \delta^2 < c < \delta \tag{8}$$

(*i*) *If*

$$C > \delta + \left(n^B - 1\right)\delta^2 - c \tag{9}$$

*then two disconnected star-shaped communities is a pairwise equilibrium network (**complete segregation**). All blues are **separated**.*

(*ii*) *If*

$$\delta - \delta^3 - c < C < \delta + \left(n^B - 1\right)\delta^2 - c \tag{10}$$

*then star-shaped communities connected through their central agents is a pairwise equilibrium network (**partial integration**). Some blues are **separated** and some are **integrated** but none has **oppositional identity**.*

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

(iii) If

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

and

$$C < \min\left\{\delta + \delta^2 - \delta^4 - \delta^5 - c, 4\left[c - \left(\delta - \delta^3\right)\right]\right\} \tag{11}$$

*then star-shaped communities where each peripheral agent has one bridge link with the other peripheral agent whereas stars have no bridge links is a pairwise equilibrium network (**partial integration**). Some blues are **separated** and some are **integrated** but none has **oppositional identity**.*

(iv) If

$$C < \delta - \delta^3 - c \tag{12}$$

*then star-shaped communities where one star is connected to the other star and all peripheral agents from both communities are connected to each other is a pairwise equilibrium network (**partial integration**). In that case, **oppositional identities** emerge in equilibrium and all blues are **integrated**.*

Figure 10 displays the different cases of Proposition 5 for $n^B = n^G = 3$.

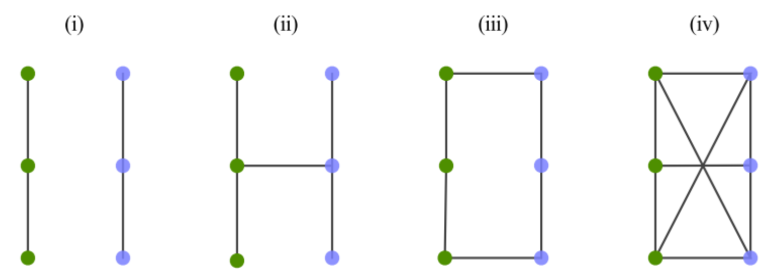


Figure 10. Different equilibrium networks when $\delta - \delta^2 < c < \delta$.

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 oppositional identities emerge in case (iv), i.e. why some blues have most of their friends who are blues (but are still *integrated*) and others have most of their friends who are greens (but are still *integrated*). In case (iv), each peripheral blue (green) 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 oppositional identities. 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 a 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 intercommunity 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 now easy to understand why a blue central individual has most of his friends who are blues. This is due to the fact that he is not attractive 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 again 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.

In terms of equilibrium utility, let us study the most interesting case, i.e. (*iv*). The utility of the peripheral individual (oppositional) is

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

while that of the center (non-oppositional) is:

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

We have

$$U_P \gtreqless U_C$$

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

As above, this condition is not incompatible with (12) and thus the oppositional individual can have a higher or lower utility than the non-oppositional one depending on the value of C as compared to δ.

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. Our analysis shows that there is a range of parameters in which two extreme outcomes, the complete network (in which all pair of agents, no mater their respective types, are connected) and a segregated network (in which only the connections among same type agents are established), are both stable networks. This is case (*ii*) in Proposition 1 where conditions $c < \delta - \delta^2$ and (4) need to hold for these two equilibria to coexist together.[22] 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 perspective. We shed here some light on this issue.

We undertake a utilitarian perspective, in which social welfare is measured by the 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 of the two networks is socially preferable.

Proposition 6 *Assume $c < \delta - \delta^2$ and (4). If*

$$n^B \left(n^G - 1\right) \left(n^B - 1\right) \leq (n-1)^2 \tag{13}$$

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*. Nevertheless, it allows us to extract some preliminary conclusions on the possible (in)effectiveness of policies that can favor socialization and thus interaction between different communities. Policies that diminish intracommunity 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 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 active community.

5 Extensions

5.1 Different externalities

We now extend our model by considering different benefits from interacting with others. Basically, if someone (whatever her type) has a link with a green (blue), she obtains a direct benefit of δ_G (δ_B). We also assume the same structure for indirect benefits. For example, if someone is connected

[22] Note that under $c < \delta - \delta^2$ and (4), no other equilibrium networks can emerge.

to a green who has a blue friend, then she gets $\delta^G + \delta^G \delta^B$. The cost structure is exactly as before and given by (2). The benefit δ^τ can be interpreted in different ways. If, for example, we think of teenagers in a school, then δ^τ could represent the human capital of individual $i(\tau)$'s parents so that being friend with someone creates positive externalities in terms of education. If, for example, we think of adults in the labor market, then δ^τ could represent the exchange of job information between two connected individuals. As stated above, strong ties are people from the same community while weak ties are those from the other communities. If greens have a better network than blues, then, as argued by Granovetter (1973, 1974), (green) weak ties are superior to (blue) strong ties for providing support in getting a job because closed networks are limited in providing information about possible jobs. In a close network, everyone knows each other, information is shared and so potential sources of information are quickly shaken down, the network quickly becomes redundant in terms of access to new information. In contrast, Granovetter stresses the strength of weak ties involving a secondary ring of acquaintances who have contacts with networks outside ego's network and therefore offer new sources of information on job opportunities.

We assume that $\delta^G > \delta^B$ so that there is a higher benefit of interacting with a green than with a blue, i.e. the direct externality green individuals exert on others is larger than the one exerted by blue individuals. In the case of teenagers, because it is well-documented that on average greens have higher human capital than greens (see e.g. Neal, 2006), then interacting with a green provides a higher benefit in terms of education for students. In the labor market interpretation, since greens have in general better information on jobs than blues (because they are more likely to be employed and the employers are more likely to be green), then the benefits to interact with greens should also be higher. Benabou (1993) has a similar assumption in his model with high and low types. There is an asymmetry between the two types in the sense that low types benefit more from high types than the reverse.

So basically, we will have the following trade off. On the one hand, blues want to interact with blues because it is less costly. On the other hand, they want to interact with greens because they obtain more direct (and indirect) benefits. For greens, it is more likely than they will mostly interact with greens since it is both less costly and leads to higher benefits.[23]

Let us focus on the case where c is low enough (case of $c < \delta - \delta^2$ in Proposition 1) which will translate here by $c < \min\left\{\delta^G - \left(\delta^G\right)^2, \delta^B - \left(\delta^B\right)^2\right\}$. To guarantee that this condition is always true, we assume:

$$c < \delta^B \left(1 - \delta^G\right) \tag{14}$$

We have shown that, when $\delta^B = \delta^G = \delta$, and $c < \delta - \delta^2$, then no individual could have an oppositional identity unless communities have different sizes such that $n^G > n^B$. Let us now show

[23] Austen-Smith and Fryer (2005) and Battu et al. (2007) also model these types of trade off in the context of black and white individuals. Our model goes further in the analysis by explicitly introducing a network formation analysis. This allows us to show that, not only direct (strong ties) matter but also indirect peer (weak ties or friends of friends) effects matter.

that one can obtain oppositional identities even if (14) holds and $n^G = n^B = n/2$ as long as $\delta^G > \delta^B$.

Consider the network described in Figure 11. There are four types of agents. From the blue population, there are two blue individuals (referred to as B_m) who are connected to all individuals in the network and therefore has n^G green friends and $n^B - 1$ blue friends. They have an oppositional identity since $n^G > n^B - 1$, meaning that they have more green than blue friends. They are also integrated since they have both green and blue friends. All the $n^B - 1$ other blue individuals (type B_1) are not connected to any green are thus separated. From the green population, $n^B - 1$ of them have two blue friends each while $n^G - \left(n^B - 1\right)$ of them have one blue friends each. The features of this particular network is somehow consistent with the friendship relationships of teenagers in the US described in Figure 2.

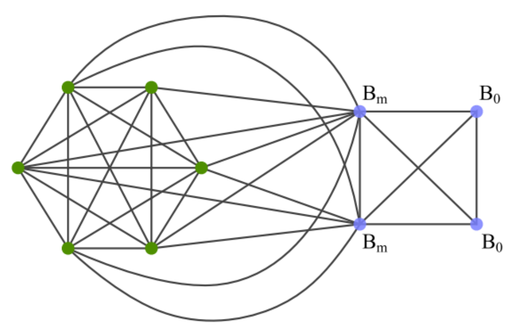


Figure 11. A network with both integrated and separated black individuals.

Let us now show under which condition the network displayed in Figure 11 can be an equilibrium network.

Proposition 7 *Assume (14). If $\delta^G = \delta^B = \delta$, the network described in Figure 11 is an equilibrium network where most blues have mostly blue friends and others (i.e. two) mostly green friends (oppositional identities) while greens have a majority of green friends. If δ^G is not too large compared to δ^B, then the network described in Figure 9 is an equilibrium network if the following condition holds:*

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

If $\delta^G >> \delta^B$, the network described in Figure 11 might not be an equilibrium network.

The intuition of this result is as follows. If $\delta^G >> \delta^B$, then greens have much less incentive to connect to oppositional blues (denoted by B_m), even if the latter have a lot of green friends. On the contrary, oppositional blues want to connect to greens because of the high externalities generated from a link with a green. In particular, a green agent might not have enough incentives to build a link with a second oppositional identity blue because the indirect externalities that she receives from the other green who is already connected to the blue community are large enough. With fixed δ^B, this can happen when the direct externalities δ^G greens exert are very large.[24] In this case, because of mutual consent, there cannot be two oppositional identity blues (i.e. type B_m). However, when δ^G is not too large compared to δ^B, then coordination problems are less sever and bridge links are easier to form. In this case, more than one blues with oppositional identity can exist in equilibrium. Observe that Proposition 7 holds if the size of the communities are the same, i.e. $n^B = n^G = n/2$. Thus, this example highlights the crucial role of coordination problems and mutual consent in friendship relationships.

5.2 Social Norms

Let us now go back to the model with the same benefits of direct interactions, δ, whatever the type, but modify the cost of creating links by taking into account *social norms*. The utility function of an individual i of type $\tau = B, G$ is now defined as:

$$u_i(g) = \sum_{j \in N \setminus \{i\}} b_{ij} - \sum_{j \in N_i(\mathbf{g})} c_{ij} \tag{15}$$

In this utility function, the benefits from connections are

$$b_{ij} = \max_{p_{ij} \in P_{ij}(g)} \omega(p_{ij}) \tag{16}$$

where $p_{ij} \in P_{ij}(g)$ is a *path* from i to j, $P_{ij}(g)$ is the set of all paths between i and j in network g, and $\omega(p_{ij})$ corresponding weights defined as follows:[25]

$$\omega(p_{ij}) = s(\left|E_{ij}^G\right|)\,\delta$$

with $E_{ij}^G = e_i^G - e_j^G$ (remember that $e_i^G \equiv n_i^G/n_i$ is the percentage of i's green friends). The function $s(\left|E_{ij}^G\right|)$ is decreasing in $\left|E_{ij}^G\right|$ and is such that $0 < s(\left|E_{ij}^G\right|) \leq 1$. In particular, $s(0) = 1$ and $s(1) = \underline{s}$, where $0 < \underline{s} < 1$.

The interpretation of (15) is as follows. The costs c_{ij} to interact with other people are still given by (2). The benefits b_{ij} are, however, different. For greens, the benefits of direct connections is δ whatever the type of the friend. For blues, the benefits of a direct connection with a green is δ while with a blue is $s(\left|E_{ij}^G\right|)\delta$. This function, which is between 0 and 1, aims at capturing the idea of *social norms* and *social norms* from the blue community. If some blues decide to have a lot of

[24] This can be seen in condition (44) in the proof of Proposition 7.

[25] For a path $p_{ij} = \langle i_0, i_1, ..., i_k \rangle$ from i to j in the network g such that $i_0 = i$, $i_k = j$, $i_p \neq i_{p+1}$, and $g_{i_p i_{p+1}} = 1$,

$$\omega(p_{ij}) = \prod_{l=0}^{l=k-1} \omega(i_l i_{l+1})$$

For $k = 1$, $p_{ij} = ij$, and $\omega(ij)$ is the unique path of length 1 between i and j.

green friends, there is a "penalty" from the blue community. The function $s(\left|E_{ij}^G\right|)$ is decreasing in the percentage of i's green friends, which means that blues obtain less and less benefits from their direct blue friends, the higher is their number of green friends. Interestingly, indirect connections are also affected by the social norms $s(.)$ for both blues and greens. This is because the social penalty reduces first the direct contact externalities and then the indirect ones.

There are studies, cited in Akerlof (1997), which illustrate the importance of social sanctions and social norms in ethnic groups. 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).[26] 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.[27]

With this new element in the utility function, there is a new force in the model: while by connecting with the other community agents become more attractive to that community, there is a cost associated with this attractiveness derived from an increased diversity in community structure. Diversified identities might dilute the positive effect of being attractive to the other community. It is interesting to note that this last effect is pairwise dependent, that is it only depends on the identity of the individual the agent is trying to connect to, while the effect of social norms is more global, since it depends on the structure of all peers identities.

Let us take the following social function

$$\begin{aligned} s(\left|E_{ij}^G\right|) &= 1-(1-\underline{s})\left|E_{ij}^G\right| \\ &= 1-(1-\underline{s})\left|E_i^G-E_j^G\right| \end{aligned} \tag{17}$$

so that $s(0)=1$ and $s(1)=\underline{s}$ where $0<\underline{s}<1$. This implies that the direct gain for a blue of interacting with another blue is: $\delta-(1-\underline{s})\left|e_i^G-e_j^G\right|\delta$, which is lower than δ, the direct gain when there were no social sanctions/norms from the blue community.

[26] 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.

[27] See also Stack (1976) for an interesting story of social sanctions/norms imposed by two systers 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.

To understand the role of sanctions/norms in the utility function, let us calculate the benefits,[28] i.e. $\sum_{j \in N \setminus \{i\}} b_{ij}$ in (15), with and without sanctions/norms for individual B_m who wants to form a link with B_0 in the network described by Figure 8. With no sanctions, the total benefits of creating this link are equal to:

$$\underbrace{\delta}_{\text{direct benefits}} + \underbrace{\left(n^B - 1\right)\delta^2}_{\text{indirect benefits}}$$

while with sanctions, there are given by:

$$\underbrace{\delta - (1 - \underline{s})\left(\frac{n^G}{n^G + 1} - 0\right)\delta}_{\text{direct benefits}} + \underbrace{\left(n^B - 1\right)\left[\delta^2 - (1 - \underline{s})\left(\frac{n^G}{n^G + 1} - 0\right)\delta^2\right]}_{\text{indirect benefits}}$$

Indeed, B_m receives a direct sanction from individual B_0, which is equal to $(1 - \underline{s})\left(\frac{n^G}{n^G+1} - 0\right)\delta$, and an indirect sanction from the $n^B - 1$ individuals who are direct friends of B_0 equal to $(1 - \underline{s})\left(\frac{n^G}{n^G+1} - 0\right)\delta^2$. The sanctions are here maximal because $\left|e_i^G - e_j^G\right|$ is the maximal value one can obtain since B_m has before the possible link with B_0 only green friends while B_0 has no green friend at all. This benefit function aims at capturing the fact that becoming totally assimilated to the green culture (i.e. having only green friends as B_m in Figure 8) has a cost if this person (B_m) wants to renew contact with her original community. There is not only a direct cost to be friend with a non-assimilated blue but also an indirect cost imposed by the whole community, i.e. the $n^B - 1$ individuals. Not surprisingly, the next result shows that the network depicted in Figure 8 will be easier to sustain in equilibrium when social sanctions/norms are introduced in the utility function.

Proposition 8 *Assume (3). If the utility is defined by (15) and social norms by (17), then if*

$$C > \left[1 - \frac{(1 - \underline{s})\, n^G}{n^G + 1}\right]\left(\frac{n^G + 1}{n^G - 1}\right)\left[\delta + \left(n^B - 1\right)\delta^2\right] - \left(\frac{n^G + 1}{n^G - 1}\right)c \tag{18}$$

the network described in Figure 8 is pairwise stable. Condition (18) is less restrictive than (7), which is the case with no social sanctions/norms.

This is an interesting result that highlights the role of social norms. Basically, when there were no social norms (Proposition 3), "attractiveness" was crucial for the result. When social norms are introduced, having green friends increase even more the distance between the *assimilated* blue B_m and the *separated* blues of type B_0. As a result, B_m does not want to be friend to a B_0 not only because she is losing her "attractiveness" with respect to the green community but also because

[28] As stated above, there are no sanctions/norms in the cost function. We could have introduced sanctions/norms in the cost function instead of the benefits, but this would not have changed the main results. It seems, however, more natural to introduce them in the benefits.

she is getting less benefits when interacting with blues. Interestingly, for the green community, B_m is still considered as a "green" person since the cost of interacting with her is still c. For the blue community, B_m is less "valuable" in terms of friend than any other blue.

6 Conclusion

In this paper, we consider social networks as the main building blocks of individual identity formation. This is a complementary view from that developed in other research such as Akerlof and Kranton (2000), where identities are sometimes interpreted as a direct choice and where it is precisely this unidimensional choice that determines socioeconomic outcomes. The choice of direct network interactions by an individual is, instead, necessarily a multidimensional and complex decision. In our case, these decentralized linking decisions are the channel determining each individual's social capital. We have modeled these decisions through a precise network structure that shapes social interactions and the exposure and assimilation of others' differences.

Identification patterns are important for individual and collective social and economic outcomes. Using Swedish data and focusing on the two-dimensional aspect of identity as defined in Figure 1, Nekby and Rödin (2009) show that what matters for labor market outcomes is strength of identification with the majority culture regardless of strength of ethnic identity. In other words, having a strong ethnic identity is not necessarily negative for the labor market if it is not associated with a rejection of the majority culture values. Using the same bidimensional measure of identity, Zimmermann et al. (2007), Constant and Zimmermann (2008), Constant et al. (2009) find, for Germany, that human capital acquired in origin countries lead to lower identification with the majority culture while education acquired post-migration, in the host country, does not affect attachment to the majority culture. Battu and Zenou (2010) find similar results for the UK while Bisin et al. (2011), studying different European countries, show that there is a penalty in the labor market for minorities with a strong identity.

We believe that our model points to an important and still understudied issue in the literature on economics and identity. In particular, our analysis has been able to mimic in equilibrium networks some characteristics of different real-world networks, such as the rise of oppositional identity patterns. In what follows, we suggest three avenues of research that seem particularly promising.

Initial exogenous differences, reflected in our model by the initial assignment of one of the two possible types, are reasonable in some setups. For example, family endows each individual with some cultural traits, such as inherited language. Yet, in other setups, we expect that both the initial identity (type) and direct connections (network) are an individual choice. This can be the case, for example, in adolescent behavior in the classroom.[29] It would be interesting to encompass

[29] See, for example, Coleman (1961), where a taxonomy of identities that adolescents adopt in US high-schools is

in a unified framework both dimensions of choice and to study the interplay of both the individual and social dimensions in the determination of identity. Presumably, in this richer framework, there might be complementarities in the final strategies of each individual in both dimensions: the choice in one dimension correlates and amplifies the choice in the other dimension.

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 an exhaustive characterization of the set of equilibrium networks. This is going to increase the already important combinatorial complexity in the analysis, that already deprives us from obtaining a full characterization of pairwise stable networks.[30]

Finally, we have not deepened another important consequence of network structure: segregation. A recent work by 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.

References

[1] Abrams D. and M. Hogg (1999), *Social Identity and Social Cognition*, Oxford: Blackwell Publishers.

[2] Ainsworth-Darnell, J.W. and D.B. Downey (1998), "Assessing the oppositional culture explanation for racial/ethnic differences in school performance," *American Sociological Review* 63, 536-553.

[3] Akerlof, G.A. (1997), "Social distance and social decisions," *Econometrica* 65, 10005-1027.

[4] Akerlof, G.A. and R.E. Kranton (2000), "Economics and identity", *Quarterly Journal of Economics* 115, 715-753.

[5] Akerlof, G.A. and R.E. Kranton (2002), "Identity and schooling: some lessons for the economics of education", *Journal of Economic Literature* 40, 1167-1201.

[6] Akerlof, G.A. and R.E. Kranton (2010), *Identity Economics: How Our Identities Shape Our Work, Wages, and Well-Being*, Princeton: Princeton University Press.

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

provided, and Akerlof and Kranton (2002), for an economic analysis of identity in schooling.

[30] For example, Kets et al. (2011) propose an interesting model that explores the manner in which 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. As a result, they develop another concept of stability where an allocation is stable in a network if no clique can gain by deviating.

[8] Austen-Smith D. and R.D. Fryer, Jr (2005), "An economic analysis of 'acting white'," *Quarterly Journal of Economics* 120, 551-583.

[9] Battu, H. McDonald, M. and Y. Zenou (2007), "Oppositional identities and the labor market," *Journal of Population Economics* 20, 643-667.

[10] Battu, H. and Y. Zenou (2010), "Oppositional identities and employment for ethnic minorities. Evidence for England," *Economic Journal* 120, F52-F71.

[11] Ben-Ner, A., McCall, B.P., Stephanec, M. and H. Wang (2009), "Identity and in-group/out-group differentiation in work and giving behaviors: Experimental evidence," *Journal of Economic Behavior and Organization* 72, 153-170.

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

[13] Berry, J.W. (1997), "Immigration, acculturation and adaptation," *Applied Psychology: An International Review* 46, 5-34.

[14] Bisin, A., Patacchini, E., Verdier, T. and Y. Zenou (2010), "Formation and persistence of oppositional identities," Unpublished manuscript, Stockholm University.

[15] Bisin, A., Patacchini, E., Verdier, T. and Y. Zenou (2011), "Ethnic identity and labor-market outcomes of immigrants in Europe," *Economic Policy*, forthcoming.

[16] Buhai, S. and M. van der Leij (2008), "A social network analysis of occupational segregation," Unpublished manuscript, Tinbergen Institute.

[17] Bramoullé, Y. and B.W. Rogers (2009), "Diversity and popularity in social networks," Unpublished manuscript, Northwestern University.

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

[19] Cabrales, A., Calvó-Armengol, A. and Y. Zenou (2011), "Social interactions and spillovers," *Games and Economic Behavior*, forthcoming.

[20] Camargo, B., Stinebrickner, R. and T. Stinebrickner (2010), "Interracial friendships in college," Working paper #15970 NBER.

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

[22] Coleman, J.S. (1961), *The Adolescent Society: The Social Life of the Teenager and Its Impact on Education*, New York: The Free Press.

[23] Constant, A., Gataullina, L. and K.F. Zimmermann (2009), "Ethnosizing immigrants," *Journal of Economic Behavior and Organization* 69, 274-287.

[24] Constant, A. and K.F. Zimmermann (2008), "Measuring ethnic identity and its impact on economic behavior," *Journal of the European Economic Association* 6, 424-433.

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

[26] Currarini, S., Jackson, M.O., and P. Pin (2009), "Identifying the roles of race-based choice and chance in high school friendship network formation," *Proceedings of the National Academy of Sciences* 107 (11), 4857-4861.

[27] Cutler, D., Glaeser, E.L. and Vigdor, J.L. (1999), "The rise and decline of the American ghetto," *Journal of Political Economy* 107, 455-506.

[28] Delpit, L. (1995), *Other People's Children: Cultural Conflict in the Classroom*, New York: The Free Press.

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

[30] Eguia, J.X. (2010), "A theory of discrimination and assimilation," Unpublished paper, New York University.

[31] Fordham S. and J.U. Ogbu (1986), "Black student'school success: Coping with the burden of acting White," *The Urban Review* 18, 176-206.

[32] Fryer, R.G. Jr. and P. Torelli (2010), "An empirical analysis of 'acting white'," *Journal of Public Economics* 94, 380-396.

[33] Golub, B. and Jackson M. (2008), "How homophily affects diffusion and learning in networks," Unpublished manuscript, Stanford University.

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

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

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

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

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

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

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

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

[42] Kets, W., Iyengar, G., Sethi, R. and S. Bowles (2011), "Inequality and network structure," *Games and Economic Behavior*, forthcoming.

[43] Kirman, A. and M. Teschl (2004), "On the emergence of economic identity," *Revue de Philosophie Economique* 9, 59-86.

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

[45] 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.

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

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

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

[49] 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.

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

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

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

[53] Neal, D. (2006), "Why has black-white skill convergence stopped?" In: E. Hanushek and F. Welch (Eds.), *Handbook of Economics of Education*, Amsterdam: Elsevier Science.

[54] Nekby, L. and M. Rödin (2009), "Acculturation identity and employment among second and middle generation immigrants," *Journal of Economic Psychology* 31, 35-50.

[55] Ogbu, J.U. (2003), *Black American Students in an Affluent Suburb: A Study of Academic Disengagement*, New Jersey: Lawrence Erlbaum.

[56] Patacchini, E. and Y. Zenou (2007), "Racial identity and education," Unpublished manuscript, Stockholm University.

[57] Phinney, J.S. (1990), "Ethnic identity in adolescents and adults: Review of research," *Psychological Bulletin* 108, 499-514.

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

[59] Ryder A.G., Alden, L.E. and D.L. Paulhus (2000), "Is acculturation unidimensional or bidimensional? A head-to-head comparison in the prediction of personality, self-identity, and adjustment," *Journal of Personality and Social Psychology* 79, 49–65.

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

[61] Selod, H. and Y. Zenou (2006), "City-structure, job search, and labour discrimination. Theory and policy implications," *Economic Journal* 116, 1057-1087.

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

[63] Wilson, W. J. (1987), *The Truly Disadvantaged: The Inner City, the Underclass, and Public Policy*, Chicago: University of Chicago Press.

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

[65] Zimmermann, L., Zimmermann, K.F. and A. Constant (2007), "Ethnic self-identification of first-generation immigrants," *International Migration Review* 41, 769–781.

APPENDIX A: PROOFS

Proof of Proposition 1.

(i) *Complete integration* between communities: There is *no* gain in utility for a green person to sever a link with a blue person, who is necessarily connected to the rest of the blue community, if:

$$\delta-\delta^2-c-\left(\frac{n^G-1}{n-2}\right)\left(\frac{n^B-1}{n-2}\right)C \tag{19}$$

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

The first term $\delta-\delta^2$ are the benefits derived from externalities of having a direct connection instead of an indirect connection with this blue person. The second term, $-c-\left(\frac{n^G-1}{n-2}\right)\left(\frac{n^B-1}{n-2}\right)C$,[31] is the cost of forming the link with this blue person. The sum of these first two terms is what intext we denoted the *connections effect*. The third term,

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

is the indirect benefit derived from the diminishing costs of maintaining a link with a blue person, once this new link is formed. Before forming the new link, the proportion of green friends among all green person's friends is $\frac{n^G-1}{n-3}$. Once the new link is created, this proportion diminishes to $\frac{n^G-1}{n-2}$, and this implies a decrease in the cost of maintaining the link with the n^B-1 blue persons from $c+\left(\frac{n^B-1}{n-2}\right)\left(\frac{n^G-1}{n-3}\right)C$ to $c+\left(\frac{n^B-1}{n-2}\right)\left(\frac{n^G-1}{n-2}\right)C$. The third term in (19) accounts for this difference in costs and is the *exposure effect* we disucssed in the main text.

The inequality (19) is equivalent to

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

which is always true since by assumption $n\geq n^B+2$, which implies the term in the right hand side of this last expression is negative, and $\delta-\delta^2-c>0$. Because of symmetry, and equivalent argument is valid to check the condition that guarantees that there is *no* gain in utility for a blue person to sever a link with a green person.

Now we have to check when there is *no* gain in utility for a green person to sever a link with another green person. The relevant inequality to make sure a green individual doesn't have such incentives is:

$$\delta-\delta^2-c \tag{20}$$

[31] The $n-2$ in the denominators come from the fact we are including the individuals involved in the link in the computation of exposure rates.

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

The strength of the connections effect is now equal to $\delta-\delta^2-c$ because we are considering an intracommunity link and, hence, its cost is equal to c. When a green agents severs a link with another green individual he is reducing his exposure to his own community and this diminishes the cost of each of the links with the other community. This exposure effect is measured by the term in the second line in (20). This term is negative, meaning that severing a link decreases the intracommunity costs. The inequality (20) is equivalent to

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

An equivalent argument shows that a blue individual doesn't want to sever a link with another blue individual if and only if

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

Since $n^G\geq n^B$, this last inequality is the condition that ensures no individual has incentives to sever a link with his own community.

(ii) Let us show that *complete segregation* between communities is an equilibrium network. 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:

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

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

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

Since $n^G\geq n^B$, and because mutual consent is necessary, then condition (4) guarantees that there is complete segregation. ∎

Proof of Proposition 2

(i) Let us denote G_1 (resp. B_1) the unique green (resp. unique blue) agent involved in the bridge link. Firstly, the agent G_1 has incentives to form a link with B_1 iff

$$\delta+\left(n^B-1\right)\delta^2-c>C \tag{21}$$

Similarly, the agent B_1 has incentives to form the link with G_1 iff

$$\delta+\left(n^G-1\right)\delta^2-c>C \tag{22}$$

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 (21) is a requirement for the network to be pairwise stable.

Under the assumption (3) we know that both communities are fully intraconnected. We have to check that no other pair of agents of different types different than G_1 and B_1 has incentives to form a link.

The green agent G_1 does not have incentives to form a link with a blue different than B_1 iff

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

which is equivalent to:

$$\left(\frac{n^G}{n^G - 2}\right)\left(\delta - \delta^2 - c\right) < C \tag{23}$$

The blue agent B_1 does not have incentives to form a link with a green different than G_1 iff

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

which is equivalent to:

$$\left(\frac{n^B}{n^B - 2}\right)\left(\delta - \delta^2 - c\right) < C \tag{24}$$

Because of mutual consent, and since $n^G \geq n^B$ only condition (23) is required.

Any other green agent different than G_1 does not have incentives to form a link with B_1 iff

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

which is equivalent to

$$\left(\frac{n^B}{n^B - 1}\right)\left[\delta + \left(n^B - 2\right)\delta^2 - \left(n^B - 1\right)\delta^3 - c\right] < C \tag{25}$$

Because of symmetry, any other blue agent different than B_1 does not have incentives to form a link with G_1 iff

$$\left(\frac{n^G}{n^G - 1}\right)\left[\delta + \left(n^G - 2\right)\delta^2 - \left(n^G - 1\right)\delta^3 - c\right] < C \tag{26}$$

Finally, any green agent other than G_1 does not have incentives to form a link with a blue other than B_1 iff

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

which is equivalent to

$$\delta + \left(n^B - 2\right)\delta^2 - \left(n^B - 1\right)\delta^3 - c < C \tag{27}$$

Any blue agent other than B_1 does not have incentives to form a link with a green different than G_1 iff

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

which is equivalent to

$$\delta + \left(n^G - 2\right)\delta^2 - \left(n^G - 1\right)\delta^3 - c < C \tag{28}$$

Because of mutual consent, only one of the conditions among (27) and (28) has to hold. Observe that if either (25) or (26) holds, then either (27) or (28) hold too.

Note that since $\delta > \delta^2 + c$ no green or blue individual different than G_1 or B_1 has incentives to sever a link with his own community. In this case, there is no exposure effect to consider because such individuals have no ties with the other community.

Gathering everything together the required conditions for the network to be pairwise stable are given by (21), (23), (25) and (26).

(ii) Firstly, in this network, a green agent with a bridge link does not have incentives to sever it iff

$$\delta - \delta^3 - c > C \tag{29}$$

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

A green agent that has already one bridge link does not have incentives to build a new one iff

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

which is equivalent to

$$\frac{n^G n^B}{\left(n^G - 1\right)\left(n^B - 1\right) - n^B}\left(\delta - \delta^2 - c\right) < C \tag{30}$$

Similarly, a blue agent does not have incentives to build a new bridge link with a green that has already a bridge link iff

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

Because of mutual consent, only (30) or (31) is needed. Since the first of these conditions is less restrictive, it suffices to ensure that this type of link is not formed.

Furthermore, 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 iff

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

which is equivalent to

$$\left(\frac{n^B}{n^B - 2}\right)\left(\delta - \delta^2 - c\right) < C \tag{32}$$

A green that is not in a bridge does not have incentives to create a link with a blue agent iff

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

Because of mutual consent, only (32) or (33) is needed. since the second of these conditions is less restrictive, it suffices to ensure that this type of link is not formed.

Since

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

condition (30) implies (32).

Finally, note that, again, as in the previous section, the green individuals that don´t have any bridge link with the other community don´t show incentives to sever any of their links with their community. Similarly, the rest of individuals green and blue individuals, that have one bridge link, don´t have incentives to sever an intracommunity link because this has no effect in the cost of their bridge link. This is because their exposure rate in the computation of the cost of the bridge link, which is equal to 1 for both sides of the link, is not affected in the case such intracommunity link is severed. This is a general rule that we use in subsequent proofs: when an individual has at most one bridge link, severing or creating intracommunity links generates no exposure effect.

Hence, gathering everything together we obtain that the two required conditions are (29) and (30).

(iii) We call oppositional blue the blue agent that has a bridge link with each of the members of the green community and we denote this agent by B_m. The oppositional blue individual B_m does not want to sever any of his bridge links iff

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

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

which is equivalent to

$$\frac{(n-2)(n-3)}{\left(n^B-1\right)\left(n^B-2\right)}\left(\delta-\delta^2-c\right)>C \tag{34}$$

A green agent does not want to sever his bridge link with the oppositional blue B_m iff

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

which is equivalent to

$$\left(\frac{n-2}{n^B-1}\right)\left[(1-\delta)\left(\delta+\left(n^B-1\right)\delta^2\right)-c\right]>C \tag{35}$$

Any of the non-oppositional blues, denoted by B_0, does not have incentives to directly connect with a green agent iff

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

which is equivalent to

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

A green agent does not have incentives to connect with a non-oppositional blue B_0 iff

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

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

which is equivalent to

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

The conditions (34) and (35) have to hold. Because of mutual consent, only one of the conditions (36) and (37) is required. Condition (36) is less restrictive than the last one, and hence, the set of required conditions are (34), (35) and (36), that can hold at the same time. ∎

Proof of Proposition 3

Consider the network described in Figure 8. There are n^G individuals who are all connected with each other. There is one blue B_m who is connected to all greens and is not connected to any other blue B_0. All the other $n^B - 1$ blues are fully connected with each other.

The blue individual B_m does not want to create a link with a blue individual B_0 iff

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

which is equivalent to

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

The blue individual B_m does not want to sever a link with a green individual iff

$$n^G\delta - n^G c > \left(n^G - 1\right)\delta + \delta^2 - \left(n^G - 1\right) c$$

which is equivalent to

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

This is always true because of assumption (3).

Observe that, because of (3), none of the blues B_0 would like to sever a link with a B_0. Because of mutual consent, when condition (7) holds, they cannot have a link with B_m because B_m does not want to.

A green will not sever a link with B_m iff

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

which always true because of (3).

Finally, because of (3), it follows that a green does not want to sever a link with another green. This is again, as in the case of the Proof of Proposition 2.*ii)* because for green agents there is no exposure effect to consider since in the network configuration we are analyzing they have only one bridge link with the blue community.

Therefore, condition (7) is enough to guarantee that the network described by Figure 6 is pairwise stable. ∎

Proof of Proposition 4

A green agent does not have incentives to sever a link with a blue agent whenever

$$\delta - \delta^3 - c \geq 0,$$

which is immediately satisfied when $c < \delta - \delta^2$.

On the other hand, a green agent does not have incentives to create a link with another green agent if

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

This condition is satisfied if C is high enough.

A similar argument holds for a blue agent. ∎

Proof of Proposition 5

(i) The center in the star formed by the green community (that we call the "green center") does not have incentives to build a link with the center in the star formed by the blue community (the blue center) iff

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

Similarly, the blue center does not have incentives to build a link with the green center iff

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

Because of mutual consent, only the first one, that coincides with (9), needs to hold

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

(ii) Firstly, if

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

then neither the green center nor the blue center has incentives to sever the bridge link that connects them.

Furthermore, none of the centers has incentives to sever a link with her own community 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.

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

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

Observe that this is satisfied by assumption.

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

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

Again, this condition is trivially satisfied.

Since these last two conditions ensure 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 with a blue peripheral iff

$$\left(1 - \delta^2\right)\delta - c < C$$

Because of symmetry, this same condition ensures that a blue peripheral agent does not have incentives to connect with a green peripheral agent. Hence, this last condition, jointly with $\delta + \left(n^B - 1\right)\delta^2 - c > C$, ensure that the network analyzed is stable.

(iii) A green peripheral agent does not have incentives to sever his bridge link with a blue peripheral iff

$$\delta - \delta^5 + \delta^2 - \delta^4 - c > C \tag{38}$$

This same condition ensures that a blue peripheral agent does not have incentives to sever his bridge link with a green peripheral agent.

A green peripheral agent does not have incentives to form a link with another green peripheral agent iff

$$\delta - \delta^3 < c \tag{39}$$

and this same condition ensures that a blue peripheral agent does not have incentives to form a link with another blue peripheral agent.

Following the same logic as in previous proofs, since each peripheral agent has one and only one bridge link, there is no exposure effect and none of these individuals has icnentives to sever the link with the center in their community.

The green center does not have incentives to form a link with the blue center iff

$$\delta - \delta^3 - c < C \tag{40}$$

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

Observe that if (39) then (40) immediately follows.

The green center does not have incentives to form a link with a blue peripheral agent iff

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

that holds by assumption. And this same condition ensures that the blue center does not have incentives to form a link with a green peripheral agent. Hence, because of mutual consent in link formation, we can ensure that no bridge link between the center of one community and a peripheral agent of the other is worth off.

A green peripheral agent does not have incentives to form a 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:

$$4\left[c - \left(\delta - \delta^3\right)\right] > C \tag{41}$$

and, once more because of symmetry, this same condition ensures that a blue peripheral does not have incentives to form a bridge link with another green peripheral agent.

Hence, the required conditions are (38), (39) and (41).

(iv) Firstly, the two centers don't have incentives to sever the bridge link that connects them iff

$$\delta - \delta^3 - c > C \tag{42}$$

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 (42) 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 individual does not have incentives to form a link with the center of the other community iff

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

$$\Leftrightarrow \delta - \delta^2 - c - \frac{1}{n^G}\frac{n^G - 1}{n^G}C + \frac{1}{n^B - 1}\frac{1}{n^G}C < 0$$

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

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 his 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.

Finally, a peripheral blue has no incentives to sever the link with the blue center if

Hence, only the first of the inequalities, $\delta - \delta^3 - c > C$, is required for that network to be 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:

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

Segregation is better if and only if:

$$\delta\leq\left[c+\frac{\left(n^G-1\right)\left(n^B-1\right)}{(n-1)^2}C\right]\left(1-\frac{n}{2n^Gn^B}\right) \tag{43}$$

It is easy to check that $1>\frac{n}{2n^Gn^B}$ 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 (43) is now given by:

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

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

$$\begin{aligned}
&\delta\left[\frac{2(n-1)^2n^Gn^B+\left(n^G-1\right)\left(n^B-1\right)\left(2n^Gn^B-n\right)}{2n^Gn^B-n}\right]\\
>\;&\left[(n-1)^2-\left(n^G-1\right)\left(n^B-1\right)\right]c+\left(n^G-1\right)\left(n^B-1\right)\left(n^B-1\right)\delta^2
\end{aligned}$$

$$\begin{aligned}
\Leftrightarrow\;&\delta\left[\frac{2n^Gn^B}{2n^Gn^B-n}-\frac{\left(n^G-1\right)\left(n^B-1\right)}{(n-1)^2}\right]-\frac{\left(n^G-1\right)\left(n^B-1\right)^2}{(n-1)^2}\delta^2\\
>\;&\left[1-\frac{\left(n^G-1\right)\left(n^B-1\right)}{(n-1)^2}\right]c
\end{aligned}$$

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

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

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

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

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

which is equivalent to (13). ∎

Proof of Proposition 7

We assume that (14) holds, so that each community is always fully connected.

A green does not want to sever a link with an oppositional blue B_m iff:

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

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

which is equivalent to:

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

A green does not want to create a link with any other blue B_0 iff:

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

which is equivalent to:

$$C > \frac{\left(n^G+1\right) n^G (n-2)}{\left(n^G-1\right)\left[n^G(n-2) - 2\left(n^B-1\right)\right]}\left[\left(1-\delta^B\right)\delta^B - c\right] \tag{45}$$

An oppositional blue B_m does not have incentives to sever any bridge link iff:

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

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

which is equivalent to

$$C < \frac{\left[\left(1-\delta^G\right)\delta^G - c\right] n^G (n-2)(n-3)}{\left(n^B-1\right)\left(n^B-2\right)\left(n^G-1\right)} \tag{46}$$

A non-oppositional blue B_0 does not have incentives to create a link with a green iff

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

which is equivalent to

$$C>\left[\left(1-\delta^{B}\right)\delta^{G}-c\right]\left(\frac{n^{G}+1}{n^{G}-1}\right) \tag{47}$$

Because of mutual consent only three conditions have to hold, that is either (44), (45), and (46) or (44), (46), and (47).

Let us start with conditions (44), (45), and (46). For these three conditions to be true, it has to be in particular that

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

which is equivalent to:

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

It is easy to see that

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

with equality if $\delta^G=\delta^B$, and that $\frac{\delta^{B}\left(1-\delta^{G}\right)-c}{\left(1-\delta^{B}\right)\delta^{B}-c}$ is decreasing in δ^G and increasing in δ^B. Furthermore, it is also easy to see that

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

As a result, firstly, we can ensure that if $\delta^G=\delta^B$ the network analyzed is pairwise stable, and by a continuity argument this network remains pairwise stable for some range of parameters in which $\delta^G>\delta^B$. Secondly, for fixed n^G and n^B, if δ^G is very large compared to δ^B, inequality (48) might not hold.

Let on now consider conditions (44), (46), and (47). For these three conditions to be true, it has to be in particular that

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

which is equivalent to

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

It is easy to see that $\frac{\delta^B(1-\delta^G)-c}{(1-\delta^B)\delta^G-c}$ is smaller or equal than 1 (with equality if $\delta^G = \delta^B$) and that it is decreasing in δ^G and increasing in δ^B. And, again, it is easy to see that

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

which ensures that, indeed, the four conditions (44), (45), (46), and (47) are compatible at the same time when $\delta^G = \delta^B$ and also, by a continuity argument, for a range of parameters with $\delta^G > \delta^B$. Anyhow, for fixed n^G and n^B, if δ^G is very large compared to δ^B, inequality (49) might neither hold. ▮

Proof of Proposition 8

Consider the network described in Figure 8. There are n^G individuals who are all connected with each other. There is one blue B_m who is connected to all greens and is not connected to any other blue B_0. All the other $n^B - 1$ blues are fully connected with each other.

The blue individual B_m does not want to create a link with a blue individual B_0 iff

$$\begin{aligned}
& n^G\delta - n^G c - n^G\left(0 \times \frac{n^G-1}{n^G}\right)C \\
> \;& n^G\delta + \delta - (1-\underline{s})\left|\frac{n^G}{n^G+1} - 0\right|\delta + \left(n^B-1\right)\delta^2\left[1-(1-\underline{s})\left|\frac{n^G}{n^G+1}-0\right|\right] \\
& -c - n^G c - n^G\left(\frac{1}{n^G+1}\frac{n^G-1}{n^G}\right)C
\end{aligned}$$

which is equivalent to

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

which is (18).

As in the proof of Proposition 3, condition (3) guarantees that no individual in this network wants to create or sever a link.

Let us now show that this condition is less restrictive that (7), i.e. the one in Proposition 3 where there are no social norms. We want to show that

$$\begin{aligned}
& \left[1-\frac{(1-\underline{s})\,n^G}{n^G+1}\right]\left(\frac{n^G+1}{n^G-1}\right)\left[\delta+\left(n^B-1\right)\delta^2\right]-\left(\frac{n^G+1}{n^G-1}\right)c \\
< \;& \frac{\left[\delta+\left(n^B-1\right)\delta^2-c\right]\left(n^G+1\right)}{n^G-1}
\end{aligned}$$

which is equivalent to

$$\underline{s} < 1$$

which is always true by definition. ▮

Proof of Proposition 9 (from Appendix B)

(i) Let us first show that *complete integration* between communities is always an equilibrium network. There is *no* gain in utility for a green person to sever a link with a blue person, who is necessarily connected to the rest of the blue community, if:

$$\delta-\delta^{2}-c-\left[k+\left(\frac{n^{G}-1}{n-2}\right)\left(\frac{n^{B}-1}{n-2}\right)\right]C \tag{50}$$

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

The first term $\delta-\delta^{2}$ are the benefits derived from externalities of having a direct connection instead of an indirect connection with this blue person. The second term, $-c-\left[k+\left(\frac{n^{G}-1}{n-2}\right)\left(\frac{n^{B}-1}{n-2}\right)\right]C$, is the cost of forming the link with this blue person. Observe that, before forming the link, the proportion of green friends among all green person's friends is $\frac{n^{G}-1}{n-2}$, while the proportion of blue friends among all blue person's friends is $\frac{n^{B}-1}{n-2}$. The third term,

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

is the indirect benefit derived from the diminishing costs of maintaining a link with a blue person, once this new link is formed. Before forming the new link, the proportion of green friends among all green person's friends is $\frac{n^{G}-1}{n-3}$. Once the new link is created, this proportion diminishes to $\frac{n^{G}-1}{n-2}$, and this implies a decrease in the cost of maintaining the link with the $n^{B}-1$ blue persons from $c+\left(\frac{n^{B}-1}{n-2}\right)\left(\frac{n^{G}-1}{n-3}\right)C$ to $c+\left(\frac{n^{B}-1}{n-2}\right)\left(\frac{n^{G}-1}{n-2}\right)C$. The third term in (50) accounts for this difference in costs.

The inequality (50) is equivalent to

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

The RHS of this inequality is negative iff:

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

A sufficient condition is thus:

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

Similarly, because of symmetry, the condition that guarantees that there is *no* gain in utility for a blue person to sever a link with a green person is

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

which is always true if (53) is satisfied. As a result, *complete integration* between communities is always an equilibrium network if (53) is satisfied.

(*ii*) Let us show that *complete segregation* between communities is an equilibrium network. 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:

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

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

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

Since $n^G \geq n^B$, and because mutual consent is necessary, then condition (4) guarantees that there is complete segregation.

(*iii*) Let us find the condition that guarantees that there are no equilibrium for which each community is not fully connected. 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 found by

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

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

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

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

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

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

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

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

Putting together (4) and this condition leads to

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

■

APPENDIX B: Oppositional identities when intercommunity costs are always larger.

In the previous section, we found that bi-partite networks were pairwise stable because there were no costs of becoming "green" for a blue person. In the equilibrium network described in Figure 8, the blue B_m "becomes" a green for other greens since the cost of interacting with her is just c. This is due to our assumption on the cost function which stipulates that the *intercommunity* cost is equal to the *intracommunity* 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 *intercommunity* cost function for $\tau(i) \neq \tau(j)$:

$$c_{ij} = c + \left(k + e_i^{\tau(i)} e_j^{\tau(j)}\right) C \tag{51}$$

where $0 < k < 1$ (we still assume that $c_{ij} = c$ if $\tau(i) = \tau(j)$). With this new intercommunity cost function, a blue person can never become totally "green" for other greens 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 9 *Assume $c < \delta - \delta^2$ and (51). If*

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

holds, then any equilibrium network is such that each community is fully connected. In particular, a bipartite network (such that the one described in figures 8 or 9) can never be an equilibrium. Furthermore, if

$$k < \frac{\left(n^B - 1\right)\left(n^G - 1\right)}{(n-2)^2} \tag{53}$$

and

$$C > \frac{\delta + \left(n^B - 1\right)\delta^2 - c}{1 + k} \tag{54}$$

hold, then both the network for which the blue and green communities are ***totally integrated*** *and the one for which the blue and green communities are* ***completely segregated*** *are equilibrium networks.*

When the intercommunity cost function is given by (51), then each community forms a complete network is C is not too large. In that case, no bipartite network can emerge. This is because nobody can now become "like" someone from the other type and, therefore, the attractiveness of having only friends from the other community is much lower. Interestingly, when k is not too large and C high enough, each individual can either have links with all individuals (including those from the other community) or only links with her own community. 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.