Ethnic Identity and Social Distance in Friendship Formation*

Joan de Marti[†] Yves Zenou[‡]

November 11, 2009

Abstract

We analyze a model of network formation with agents that belong to different communities and an endogenous cost structure. Both individual benefits and costs depend on direct as well as indirect connections. Benefits of an indirect connection decrease with distance in the network, while the cost of a link depends on the type of agents involved 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. 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 that rely on individual positions in the network. We derive a number of results with regard to equilibrium networks: (i) socialization among the same type of agents might be weak even if the within-type link cost is very low; (ii) oppositional identity patterns can arise for a wide range of parameters; (iii) integrated networks can be socially preferable to segregated networks.

Keywords: Network formation, identity, bridges, structural holes, ethnic minorities, social norms.

JEL Classification: A14, 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, Spain. 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 focus on *ethnic identity* and analyze the friendship formation process between individuals of different ethnic groups.

Part of the literature on ethnic identity has visualized this concept 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). Studies in the US (but also in the UK) have found, for example, 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; Fryer and Torelli, 2005; Selod and Zenou, 2006; Battu et al., 2007).[2]

There is a literature in psychology (see, in particular, Berry, 1997; Phinney, 1990; Ryder et al., 2000) that proposes a broader concept of self-identification in a *two-dimensional* framework, where identifications with two different cultures are not necessary mutual exclusive. In fact, Berry (1997) identifies 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.

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

[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.

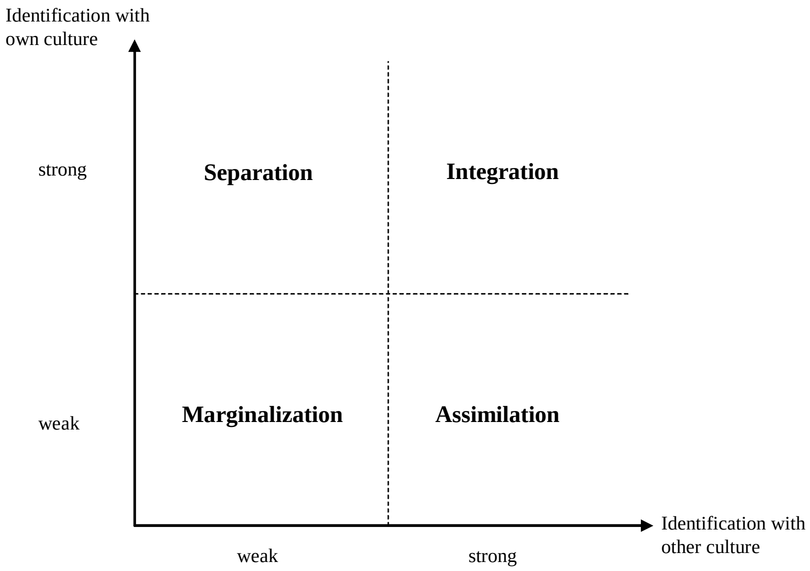


Figure 1. Different identifications for ethnic minorities

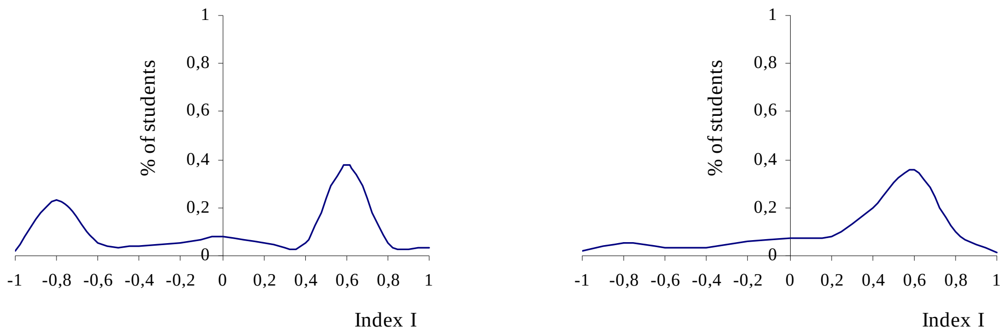
As it can be seen from Figure 1, individuals who are integrated have not only a strong identification to the majority culture but also to their own culture. Observe that the previous definition of an oppositional identity corresponds to either a separated or an assimilated individual in Figure 1.[3]

There are some empirical studies in the US using both the unidimensional and bidimensional definition of identity choices. 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.[4] 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

[3]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 (2009) find similar results for the UK.

[4]See also Currarini et al. (2007) and Pin et al. (2008) for models that have used Coleman's homophily index to show empirically how adolescents in the United States choose their friends according to 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.[5]

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.

To the best of our knowledge, there are no theoretical models explaining both the bidimensional choices of ethnic identity (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.

Model and Results

We consider a finite population of individuals composed by two different communities. These two communities can represent ethnic, racial or social groups. The composition of social relationships

[5]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.

is represented by a network. 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.[6] 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.) and it is modelled by means of a decay factor. 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.

Ethnic identity is usually characterized by a language and other distinctive cultural traits. When two individuals of different ethnic groups interact, they may experience a disutility due to the attachment to their original culture. This discomfort can, however, be mitigated if individuals are frequently exposed to the other community. Indeed, when someone spends time interacting with people from the other community, she can learn the codes and norms (prescriptions) that govern their social interactions. This is precisely the starting point of our analysis: 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 the 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.

In this respect, social distance[7] 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 while the final total cost is a function of the structure of the network.

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. Its main virtue is that it takes into account 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 severe any of her links, and no pair of agents who are not connected have incentives to build a new link.

In this context, when *intracommunity* linking costs are low, we show that oppositional identities can only 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. Observe, however, that in several equilibrium configurations *bridge links* (i.e. links that connect both communities) prevail. Even if

[6] See Goyal (2007) and Jackson (2008) for overviews of the growing literature on social and economic networks.
[7] See, for example, Akerlof (1997).

those bridge links can be quite costly for the agents involved, it gives 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. We can also determine under which condition totally *assimilated* and *separated* ethnic minorities (Figures 1 and 2) emerge in equilibrium as well as "extreme" networks (i.e. bipartite networks) when whites have only black friends and blacks only white friends. This last feature is not possible in the standard connections model. When social costs become higher, more bridge links and more interactions between communities emerge and the features described in Figures 1 and 2 can be explained even better.

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 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.[8] 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. We show that both situations facilitate the adoption of oppositional identities.

Related Literature

Besides the aforementioned literature on identity, our work relates to other works. There are indeed some recent papers that analyze homophily in social networks. 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. A paper which is closer to ours is Currarini et al. (2009). They develop a matching model with a population formed by communities of different sizes. The paper is able to replicate a number of observations from real-world data related to homophilous behavior. Our model and conclusions differ in a number of directions. First, we have an explicit model of network formation where network structure matters. Second, in Currarini et al. (2009), in equilibrium, individuals' behavior are totally homogeneous among the same group of agents. Thus, while their model replicates with high precision some of the characteristics observed in the AddHealth data, it says anything about the different identity patterns observed in Figure 2. Instead, in our model, we can observe *heterogenous behavior in equilibrium* where, for example, a fraction of individuals of a given community show oppositional identities.

Using the same dataset, Fryer and Torelli (2005) show that there are large racial differences in the relationship between students' popularity and academic achievement in high schools. Among whites, higher grades yield higher popularity. For Blacks, higher achievement is associated with

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

modestly higher popularity. A black student with a 4.0 has, on average, 1.5 fewer same-race friends than a white student with a 4.0. This is consistent with the "acting white" phenomenon and oppositional identity behavior, and it suggests that identity choices are correlated with individual socioeconomic outcomes.

Our model also contributes to the literature on strategic network formation. There are few models of network formation with differentiated communities. Jackson and Rogers (2005) extend the Jackson and Wolinsky (1996)'s connection model by including two communities. However, in their analysis, 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. Similarly, Galeotti (2006) and Galeotti et al. (2006) extend the Bala and Goyal (2000)'s connection model by introducing agents' heterogeneity but with the same cost structure as in Jackson and Rogers (2005). They focus on *directed networks* where coordination problems are not an issue given that there is no need for mutual consent. Johnson and Gilles (2000) also extend the Jackson and Wolinsky (1996)'s connection model to take into account the geographical space (modelled as individuals in a line) where people live. As a result, the cost of creating a link is proportional to the geographical distance between two individuals. Hence, 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.[9]

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 own community, can lead to large differences in terms of location decision. Indeed, his results suggest that total segregation persists even if most of the population is tolerant about heterogeneous neighborhood composition. 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). 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, racial preferences in our setup are not homogenous and are not the unique source of utility: depending on the location in the network, the same individual can have different preferences for two different agents of another race. 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. A potential partner is more attractive if she does not share most of the individual actual social capital, i.e. they have different friends and the new link gives access to other distant individuals in the population. The composition of the social capital, understood as

[9] In a very different framework, Bramoullé and Kranton (2007, 2008) analyze risk sharing decisions between members of same and different communities based on endogenous linking decisions.

the set of direct friends, of a potential partner, together with own's social capital, determines how costly it is to form a new link.

Our main contributions

We show that ex ante identical individuals only differing by their attachment to a group may end up with very different friendship patterns. 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 also 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 oppositional identities. 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. 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. Other papers such as, for example, Currarini et al. (2009) and Bramoullé and Rogers (2009), also generate asymmetries in friendships but the mechanism is not completely endogenous since there are exogenous parameters generating the results of inbreeding biases.

Another interesting and new aspect of our model is that 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 emerges. This is why it is extremely hard, if not impossible, to provide a full-fledged characterization of all possible stable networks.[10] Nevertheless, we are able to find a set of different network arrangements that are stable and display segregation and oppositional identity patterns. We do not pretend that segregation and oppositional identity patterns are the norm but we show that they can emerge in some circumstances as the result of a decentralized process of socialization. Naturally, as it is common in other models with multiple equilibria, there is no obvious way to select one equilibrium over another. The combinatorial complexity of the problem in hand is the price to pay to achieve results about a subset of network structures that arise in equilibrium.

[10] 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 our model. Indeed, 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 Network notations

There is a finite population of individuals denoted by $N = \{1, \ldots, n\}$. Each agent of the population belongs exclusively to one of the two different types, B or W. This initial endowment of each individual can be interpreted as the identity inherited from her family. The type of individual i is denoted by $\tau(i)$. Hence, an agent is of type W if and only if $\tau(i) = W$. Let n^B denote the number of B individuals in the population. Similarly, let n^W denote the number of W individuals in the population. We have that $n = n^B + n^W$. We assume, without loss of generality, that $n^B \leq n^W$.

Agents are 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 $N_i^{\tau(i)}(g) = \{j \neq i, \tau(i) = \tau(j) \mid g_{ij} = 1\}$, which is of size $n_i^{\tau(i)}(g)$.

We present some examples of network configurations. The *circle* is such that each agent has two direct contacts. The *star-shaped network* is when one central agent 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.

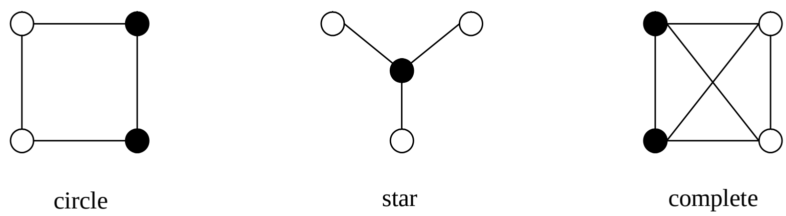


Figure 3. Circle, star and complete networks with four individuals

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. Black nodes refer to type$-B$ individuals while white nodes indicate type$-W$ 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.[11] 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

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

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

[11] 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.

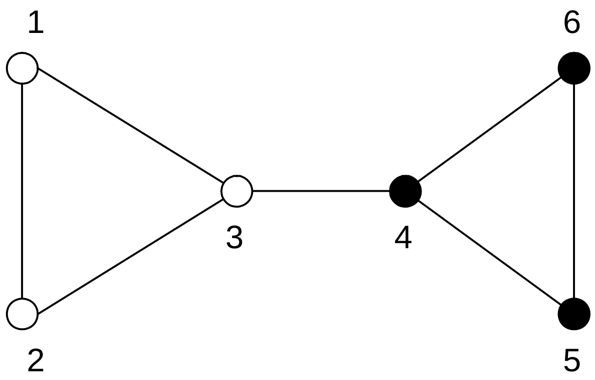


Figure 4. A bridge network

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. We assume that $c_{ij} = c$ if $\tau(i) = \tau(j)$ and

$$c_{ij} = c + q_i^{\tau(i)} q_j^{\tau(j)} C \tag{2}$$

otherwise, where

$$q_i^{\tau(i)} = \frac{n_i^{\tau(i)}}{n_i}$$

is the percentage of same-race friends of individual i ($n_i^{\tau(i)}$ is the number of i's same-race friends while n_i is the total number of i's friends independently of their type), and

$$q_j^{\tau(j)} = \frac{n_j^{\tau(j)}}{n_j}$$

is the percentage of same-race friends of individual j ($n_j^{\tau(j)}$ is the number of j's same-race friends while n_j is the total number of j's friends independently of their type)

There are thus different costs, depending with whom a connection is made. First, it is assumed that $C > 0$, which means that it is always more costly to form a friendship relationship with someone from a different race (the cost is given by (2)) than with someone from the same race (the cost is c).

Second, if an individual i of type τ forms a friendship relationship with an individual j of type τ', with $\tau \neq \tau'$ (i.e. interracial friendship formation), then, because of cultural differences and prejudices, what matters is the choice of same-race friends. If, for example, a white person has only white friends, then it will be difficult for her to interact with a black person, especially if the latter has mostly black friends. There are different cultures, norms and habits between races so that frictions are higher the more different people are. What we have in mind here is that individuals are born with a certain type τ (black or white) 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 races of types.[12] But then people make choices in terms of friendships, which could be interpreted in terms of identity. These choices can increase or decrease the original gap between individuals. If someone who is born black chooses to have only black friends (this is an identity choice), then it will be quite difficult for her to interact with a white person. However, the more similar the choices are, the easier is to interact with someone from a different type or race. Observe that we allow that friend choices can totally erase the initial gap between blacks and whites. Indeed, if at least one individual (i or j) has no friends of the same type (i.e. $q_i^{\tau(i)} = 0$ or $q_j^{\tau(j)} = 0$), then it is equally costly for them to interact with each other than with someone of the same race (i.e. the cost is c in both cases).[13] Formally, the interracial cost c_{ij} is described by(2) and it varies between c (when at least one of them has only friends of the other race, i.e. either $q_i^{\tau(i)} = 0$, $q_j^{\tau(j)} = 0$, or $q_i^{\tau(i)} = q_j^{\tau(j)} = 0$) and $c + C$ (when both have only friends of the same race, i.e. $q_i^{\tau(i)} = q_j^{\tau(j)} = 1$).[14]

To illustrate our cost function (2), consider again the network described in Figure 4 and assume that individuals 1, 2, and 3 are whites while individuals 4, 5, and 6 are blacks. 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 (White) to 4 (Black) is:

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

since $n_3^{\tau(3)} = n_4^{\tau(4)} = 2$ (number of same-race friends of 3 and 4, respectively) and $n_3 = n_4 = 2$ (total number of 3's and 4's friends independently of race),[15] which implies that $q_3^{\tau(3)} = q_4^{\tau(4)} = 1$.

[12] 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.

[13] In Section 3.1.2 below, we investigate a different cost function where the interracial cost is *not* anymore equal to the intraracial cost as soon as one of the persons involved in a relationship has no friends of the same race.

[14] The idea of looking at same-race friends as an indicator of identity is not new, at least in sociology. As mentioned in the introduction, Coleman (1958) has proposed the *homophily* index:

$$H_i = \frac{q_i^{\tau(i)} - p^{\tau(i)}}{1 - p^{\tau(i)}}$$

where $p^{\tau(i)} \equiv n^{\tau(i)}/n$ is the percentage of students of race τ in the society. When $q_j^{\tau(j)} > p^{\tau(j)}$, there is *inbreeding homophily* in the sense that individuals of a given race prefer to interact with similar people. See Curriarini et al. (2009) for a detailed analysis on this issue.

[15] It is important to observe that, when individual 3 considers the possibility of creating a link with individual 4,

Consider now the same network but with individual 4 already having a link with 2. In that case, the cost of connecting 3 (White) to 4 (Black) is:

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

since $q_3^{\tau(3)} = 1$ but $q_4^{\tau(4)} = 2/3$. Indeed, it is now less costly for individual 3 (White) to be friend to individual 4 (Black), because the latter has already a white friend.

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.[16] 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 whenever it is pairwise stable.

3 Stable networks

3.1 Low intra-community costs

We start our analysis with low intra-community costs c. In particular, we start with the case: $c < \delta - \delta^2$, since, in this cost range, any two agents who are not directly connected may benefit from forming a link. This means, in particular, that if there were only one community (i.e. only one

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

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

type of individuals), then the complete network would be the unique equilibrium network (as in the connections model of Jackson and Wolinsky, 1996). Because, we have two different communities and different cost structures, this is not true anymore. In this section, we focus on networks where each community is fully intra-connected, i.e. forms a complete network. In the next section, we will consider the case of non-fully intra-connected 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. We have a first result:[17]

Proposition 1 *Assume*

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

Then,

(i) Whatever the value of C, the network such that the black and the white communities are ***totally integrated*** *is always an equilibrium.*

(ii) If

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

holds, then the network for which the black and the white communities are ***completely segregated*** *and each community is fully connected is also an equilibrium.*

(iii) If

$$C < \delta + \left(n^B - 1\right)\delta^2 - c \tag{5}$$

holds, then all equilibrium networks are ***partially integrated*** *and each community is fully connected.*

When within-community costs are low enough (i.e. $c < \delta - \delta^2$), all individuals within and between communities are likely to be connected with each other, and, this is why the network where both communities are totally integrated is always an equilibrium. Indeed, once this network is formed, the only possible deviation is for one person to delete a link (since nobody can form a new link). Because the gain of severing a ling is $\delta - \delta^2$ while the cost is much higher (the person who deletes a link becomes less attractive to all individuals from the other community with whom she is linked to), no deviation is profitable and this network is pairwise stable. There are other equilibria as described by cases (*ii*) and (*iii*) in Proposition 1 and they mostly depend on the between-community cost C and the benefits δ. When C is high enough compared to δ (i.e. condition (4) holds), then no individual wants to be friend with someone from the other community. If we use the two-dimensional definition of identity, illustrated in Figure 1, the blacks are here *separated*. This could be the

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

case where the two populations are totally physically separated (i.e. spatially segregated) so that interactions are very costly (because, for example, of commuting costs, prejudices, etc.). Observe that, a natural condition for full segregation is $C > \delta + \delta^2(n-1)$ (i.e. the between-community cost is higher than the cost of having a link between two members of the different communities). However, because $c < \delta - \delta^2$ (see (3)), which implies that $\delta/2 + \delta^2 (n/2 - 1)/2 + c/2 < \delta + \delta^2(n-1)$, our condition (4) is weaker and therefore encompasses this case.

When C decreases relative to δ, then some individuals start forming links with the other community. There are now bridges between the two communities and partial integration prevails. This implies that some blacks are *integrated* since they have friends from both communities.

To understand our results, let us highlight 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.

(2) 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.

(3) 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.

Equilibrium networks are those that correctly balance these three forces at the individual level.

When C decreases, individuals start to form bridge links, which make them more attractive to the other community members, who, in turn, form bridge links, etc.

Let us investigate in more details the partially-integrated case (iii). Define

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

The following proposition fully characterizes case (ii):

Proposition 2 *Assume (3) and (5).*

(i) *If the following condition holds:*

$$C > \max\left\{\frac{n^W\left(\delta - \delta^2 - c\right)}{n^W - 2}, \Phi^{\tau}\left(n^W, \delta, c\right), \Phi^{\tau}\left(n^B, \delta, c\right)\right\} \tag{6}$$

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

(ii) *If*

$$\frac{n^W n^B\left(\delta - \delta^2 - c\right)}{\left(n^W - 1\right)\left(n^B - 1\right) - n^B} < C < \delta - \delta^3 - c \tag{7}$$

holds, then the network where both communities are fully intraconnected and each white and each black individual has one, and only one, bridge link is an equilibrium (Figure 6).

Basically, in case (*iii*) of Proposition 1, no individual has two bridge links and equilibrium networks are characterized by either only one bridge link (Figure 5) or the number of bridges connecting the two communities is equal to n^B (Figure 6). Some integration between blacks and whites is taking place. Let give some intuition of these results. Because we are in the partially-integrated case of Proposition 1, both conditions (3) and (5) have to hold. Indeed, C has to be low enough for (5) to hold and, since each community is fully connected (this is due to (3)), each individual prefers to be friend with someone from the same type than from the other type. As a result, because the cost of a bridging a link is quite high, no individual, who has already a bridge link, wants to form an additional one (since the benefit is $\delta - \delta^2$ while the cost is more than C).

In case (*i*) of Proposition 2, illustrated by Figure 5 (for $n^W = 6$ and $n^B = 4$), C has to be large enough because no individuals than those who already have a bridge link want to form a link with someone from the other community. Indeed, because each intra-community network is complete and there is already a bridge link, then every person is at most at distance 2 from any person from the other community. So when (6) holds, there is no incentive to form more than one bridge link.

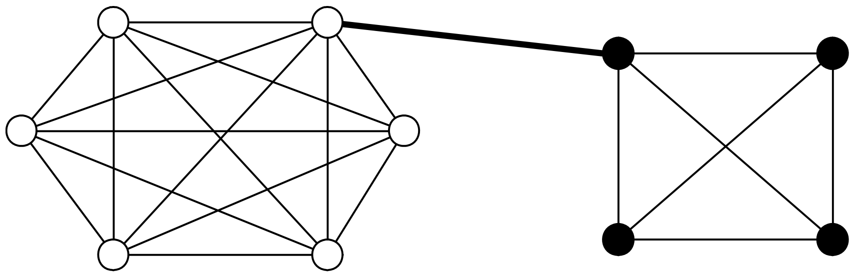


Figure 5. Equilibrium network when condition (6) holds

In case (*ii*) of Proposition 2, illustrated by Figure 6 (for $n^W = 6$ and $n^B = 4$), C cannot be too large and this is why new bridge links are formed (compared to the previous case). Even though, every person is still at most at distance 2 from any person from the other community, it becomes beneficial to form new bridge links because of a lower C.

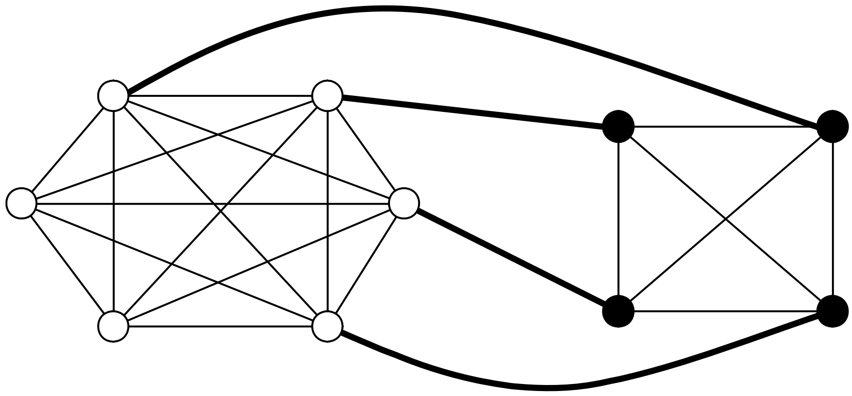


Figure 6. Equilibrium network when condition (7) holds.

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). 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).

Let us now investigate the issue of oppositional identities.

Proposition 3 *Assume (3). Then, in the totally integrated equilibrium (case* (i) *of Proposition 1), whites have most of their friends who are whites but blacks can have more white than black friends (oppositional identities).*

Indeed, from Propositions 1 and 2, when (3) holds, one can see that either individuals have no bridge link (case (ii) in Proposition 1) or there is at most one bridge link (case (iii) in Proposition 1 and cases (i) and (ii) in Proposition 2). This implies that no oppositional identity behavior can emerge in equilibrium. There are two main reasons for this. First, the cost of interacting with same-race friends is very low ($c < \delta - \delta^2$) so it is always beneficial to be friend with all individuals of the same type. This, in particular, implies that the percentage of same-race friends $q_i^{\tau(i)}$ each individual i has is very high. Second, because $q_i^{\tau(i)}$ is high, the cost of interacting with the other community is also high and thus deter people to have multiple bridge links.

In case (i) of Proposition 1, however, we have the extreme case of *full integration*, i.e. all blacks are integrated. Oppositional identities arise only because the black community is smaller in size than the white community. Indeed, each black has $n^B - 1$ black friends and n^W white friends and since $n^W > n^B$, blacks have more white friends than black friends whereas the reverse is not true. This is an extreme case since all blacks are oppositional. Observe that there are reinforcing effects. On the one hand, everybody is connected to everybody whatever the race. On the other hand, because each individual has a lot of friends from the other community, then they are both "attractive" for each other, which induce people to create even more bridge links. This also resolves the coordination problem of mutual agreements of forming links.

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 black individuals who have different positions in the network may have different incentives to form a link with a white person so that, in equilibrium, only one of them will find it beneficial to form a bridge link.

Case (i) in Proposition 1 is an extreme case of oppositional identities. Let us show now that oppositional identities can arise in less extreme case, and in particular some blacks will have a majority of black friends while other will have a majority of white friends. Because we now focus on oppositional identities and assimilation issues, type$-B$ and type$-W$ individuals will be interpreted as black and white individuals. Our analysis is, of course, valid for other interpretations in terms of types or traits.

Proposition 4 *Assume (3). If*

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

holds, then the network in which both communities are fully intraconnected and only one black agent is connected to all the agents of the other community is an equilibrium (Figure 7). There are thus oppositional identities since most blacks have mostly black friends and others (i.e. one) mostly white friends while whites have a majority of white friends.

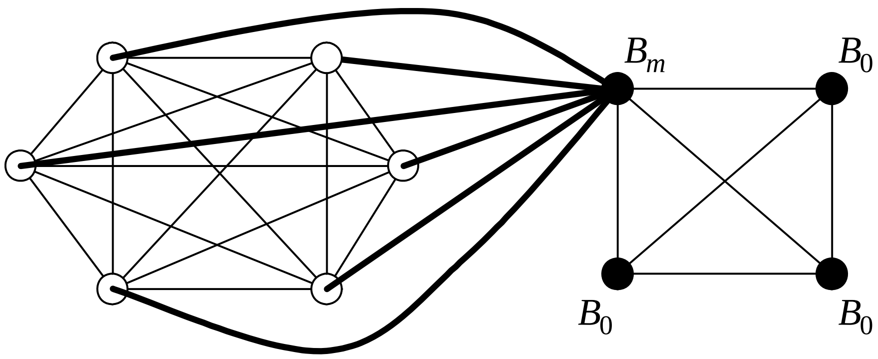


Figure 7. Oppositional identities when $c<\delta-\delta^2$

This case of oppositional identity is less extreme than that of case (i) in Proposition 1 because not all black individuals are friends to all whites and it is not the size of minority group that leads to the results. Here only one black person is oppositional (individual B_m) but also *integrated* (see Figure 1) while all other blacks are *isolated* (they have no white friends). Observe that this result still holds if the size of the communities are the same, i.e. $n^W = n^B = n/2$ (see the proof in the Appendix). So, contrary to case (i) in Proposition 1, what drives the result is not the fact that community sizes are different but rather the fact that, once one black person starts to have multiple links, she becomes "cheaper" for whites to interact with her. It is really the coordination problem due to the indirect nature of the network that is at the heart of this result. Even if C is relatively low, it has to be beneficial for both individuals to form a link with each other. This is possible only if the percentage of same-race friends $q_i^{\tau(i)}$ is low enough.

Observe also that it is not always true that oppositional individuals obtain a higher equilibrium utility than non-oppositional blacks. Take, for example, Proposition 4 (Figure 7). The equilibrium utility of the oppositional black B_m is

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

while the utility of non-oppositional blacks is:

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

So we have

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

This inequality is not incompatible with the condition given in Proposition 4, 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 whites 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 too costly to be friend with whites.

3.1.1 Assimilation Patterns

Let us still assume very low intra-community costs, i.e. $c < \delta - \delta^2$ and show that, contrary to the standard connections model, communities can be not fully connected and that oppositional identities and integration can arise because of "attractiveness" reason and coordination problems.

Proposition 5 *Assume (3). If*

$$C > \frac{\left[\delta + \left(n^B - 1\right)\delta^2 - c\right]\left(n^W + 1\right)}{n^W - 1} \tag{8}$$

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

This is an interesting result because, when (3) holds, in the connection model (Jackson and Wolinsky, 1996), all agents (of the same type) have links with each other. This is not true in this model because, by connecting to someone of the same race, a person changes her attractiveness to the other community.

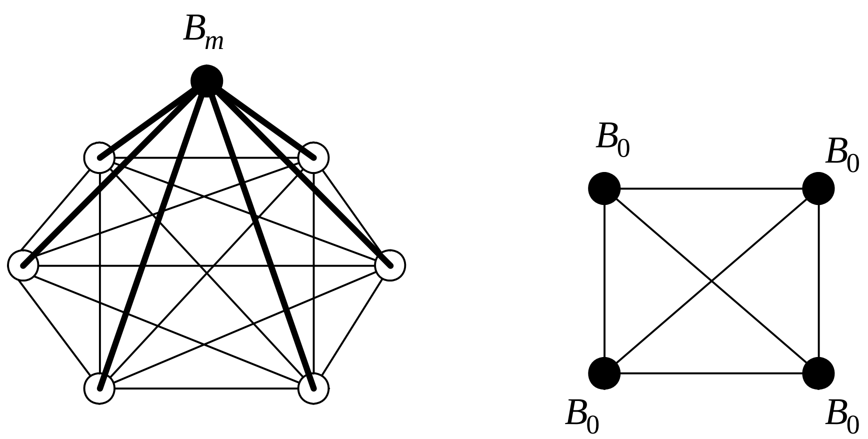


Figure 8. Oppositional identities with non-fully intraconnected communities

More generally, this result highlights the role of both "attractiveness" to the other community and of mutual consent. Indeed when (3) holds, i.e. $c < \delta - \delta^2$, then, whatever the value of C, any black B_0 would like to form a link with B_m. However, when $C > \left[\delta + \left(n^B - 1\right)\delta^2 - c\right]\left(n^W + 1\right) / \left(n^W - 1\right)$, the black individual B_m does not want to create a link with a B_0 because the gain of this link, $\delta + \left(n^B - 1\right)\delta^2$ is lower than the cost of this link, $c + n^W\left(\frac{1}{n^W+1}\frac{n^W-1}{n^W}\right)C$. Indeed, when this link is not created, B_m has a value of same-race friends $q^{\tau}_{B_m} = 0$, which makes her very attractive for the white community since, for a white person, the cost of being friend with B_m is equal to the cost of being friend with a white, i.e. c. However, when B_m is creating a link with a black B_0, her value of same-race friends $q^{\tau}_{B_m}$ increases from 0 to $1/\left(n^W + 1\right)$. With this extra link, for a white, the cost of being friend with B_m switches from c to $c + \left(\frac{1}{n^W+1}\frac{n^W-1}{n^W}\right)C$, which is much higher than c, especially when C is high enough. Since B_m has n^W friends, by not being friend with a black B_0, she is saving $c + n^W\left(\frac{1}{n^W+1}\frac{n^W-1}{n^W}\right)C$. There is thus an asymmetry in the friendship decision because B_0 wants to be friend with B_m but B_m does not. More generally, we obtain this result because, as long as a black person such as B_m is totally isolated from her community, she is basically becoming a "white" and interacting with her is as costly as interacting with a white for the white community.

This is an interesting result, which highlights the fact that assimilating to the majority culture (see Figure 1) makes it difficult for a black person to interact with her own group. In Section 5, we further investigate this case by looking at social norms and sanctions where assimilation to the white culture leads to a rejection from the black community.

Observe that in this network (Figure 8), the black oppositional B_m has always a higher utility than any other non-oppositional black B_0 since $(\delta - c)\, n^W > (\delta - c)\, n^B$.

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

Proposition 6 *Assume (3). If C is sufficiently large, the bipartite network in which all white agents are connected to all black agents, and all black agents are connected to all white 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.

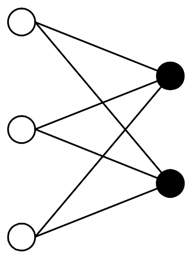


Figure 9. Bipartite Network with $n^W = 3$ and $n^B = 2$.

This network can be sustained in equilibrium because the cost of an intercommunity link reduces to the cost of an intracommunity link whenever the two agents involved do not have any link with their own community. 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 white agents are "becoming" blacks while all black agents are "becoming" whites.

3.1.2 Oppositional identities when it is costly to have only white friends

In the previous section, we found that bi-partite networks were pairwise stable because there were no costs of becoming "white" for a black person. In the equilibrium network described in Figure

8, the black B_m "becomes" a white for other whites since the cost of interacting with her is just c. This is due to our assumption on the cost function which stipulates that the *interracial* cost is equal to the *intraracial* cost as soon as one of the persons involved in the relationship has no friends of the same race. In the present section, we relax this assumption and assume instead the following *interracial* cost function for $\tau(i) \neq \tau(j)$:

$$c_{ij} = c + \left(k + q_i^{\tau(i)} q_j^{\tau(j)}\right) C \tag{9}$$

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

Proposition 7 *Assume $c < \delta - \delta^2$ and (9). If*

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

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

$$k < \frac{\left(n^B - 1\right)\left(n^W - 1\right)}{(n-2)^2} \tag{11}$$

and

$$C > \frac{\delta + \left(n^B - 1\right)\delta^2 - c}{1 + k} \tag{12}$$

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

When the interracial cost function is given by (9), 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.

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.[18] 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 8 *Assume that*

$$\delta - \delta^2 < c < \delta \tag{13}$$

(i) If

$$C > \delta + \left(n^B - 1\right)\delta^2 - c \tag{14}$$

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

(ii) If

$$\delta - \delta^3 - c < C < \delta + \left(n^B - 1\right)\delta^2 - c \tag{15}$$

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

(iii) If

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

and

$$C < \min\left\{\delta + \delta^2 - \delta^4 - \delta^5 - c, 4\left[c - \left(\delta - \delta^3\right)\right]\right\} \tag{16}$$

*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 blacks are **separated** and some are **integrated** but none has **oppositional identity**.*

(iv) If

$$C < \delta - \delta^3 - c \tag{17}$$

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

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

***(partial integration)**. In that case, **oppositional identities** emerge in equilibrium and all blacks are **integrated**.*

Figure 10 displays the different cases of Proposition 8 for $n^B = n^W = 3$.

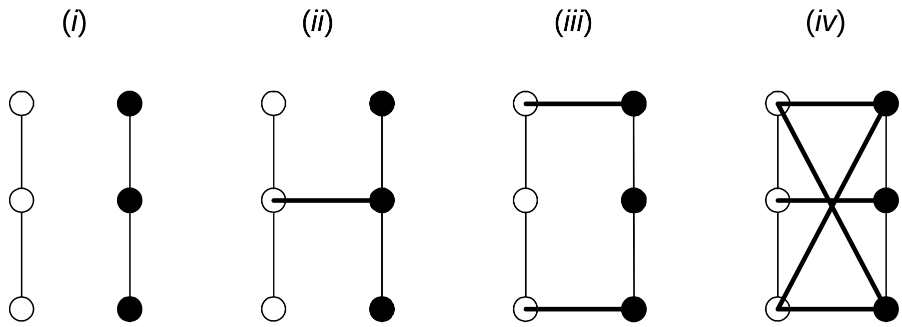


Figure 10. Different equilibrium networks when $\delta - \delta^2 < c < \delta$.

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 blacks have most of their friends who are blacks (but are still *integrated*) and others have most of their friends who are whites (but are still *integrated*). In case (*iv*), each peripheral black (white) has one black (white) friend (the central agent) and $n^W - 1$ ($n^B - 1$) white (black) friends so that their common same-race friend percentage is $q_i^{B(i)} = 1/(n^W)$ ($q_i^{W(i)} = 1/(n^B)$). This is quite small, especially when the size of the population of each community is large. As a result, each black (white) peripheral individual displays a high taste for other-race friends, which makes them very attractive. On the contrary, the black (white) central agent has one white (black) friend and $n^B - 1$ ($n^W - 1$) black (white) friends so that $q_i^{B(i)} = (n^B - 1)/n^B$ ($q_i^{W(i)} = (n^W - 1)/n^W$). 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 blacks. First, peripheral blacks 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 blacks do not want to severe a link with one of the $n^W - 1$ peripheral whites because the latter are all very attractive. Finally, peripheral blacks do not want to create a link with a central white person because she is not very attractive due to her high interracial costs and they can reach him from a peripheral white (distance 2) and obtains δ^2. This is why peripheral blacks have most of their friends who are whites. It is now easy to understand why a black central individual has most of his friends who are blacks. This is due to the fact that he is not attractive to the peripheral whites.

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 $n^B = n^W = n/2$. 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

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

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

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

We have

$$U_P \gtreqless U_C$$

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

As above, this condition is not incompatible with (17) 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 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 white (black), she obtains a direct benefit of δ_W (δ_B). We also assume the same structure for indirect benefits. For example, if someone is connected to a white who has a black friend, then she gets $\delta^W + \delta^W \delta^B$. 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 whites have a better network than blacks, then, as argued by Granovetter (1973, 1974), (white) weak ties are superior to (black) 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 $\delta^W > \delta^B$ so that there is a higher benefit of interacting with a white than with a black, i.e. the direct externality white individuals exert on others is larger than the one exerted by black individuals. In the case of teenagers, because it is well-documented that on average whites have higher human capital than whites (see e.g. Neal, 2006), then interacting with a white provides a higher benefit in terms of education for students. In the labor market interpretation, since whites have in general better information on jobs than blacks (because they are more likely to be employed and the employers are more likely to be white), then the benefits to interact with whites 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, blacks want to interact with blacks because it is less costly. On the other hand, they want to interact with whites because they obtain more direct (and indirect) benefits. For whites, it is more likely than they will mostly interact with whites since it is both less costly and leads to higher benefits.[19]

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 $c < \min\left\{\delta^W - \left(\delta^W\right)^2, \delta^B - \left(\delta^B\right)^2\right\}$. To guarantee that this condition is always true, we assume:

$$c < \delta^B \left(1 - \delta^W\right) \tag{18}$$

We have shown that, when $\delta^B = \delta^W = \delta$, and $c < \delta - \delta^2$, then no individual could have an oppositional identity unless communities have different sizes such that $n^W > n^B$. Let us now show that one can obtain oppositional identities even if (18) holds and $n^W = n^B = n/2$ as long as $\delta^W > \delta^B$.

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

[19] 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.

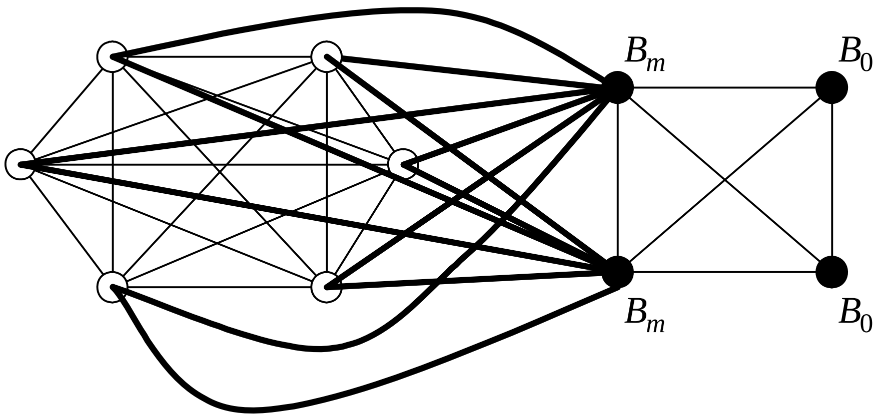


Figure 11. A network with both integrated and separated black individuals

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

Proposition 9 *Assume (18). If $\delta^W = \delta^B = \delta$, the network described in Figure 11 is an equilibrium network where most blacks have mostly black friends and others (i.e. two) mostly white friends (oppositional identities) while whites have a majority of white friends. If δ^W is not too large compared to δ^B, then the network described in Figure 9 is an equilibrium network if the following condition holds:*

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

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

The intuition of this result is as follows. If $\delta^W >> \delta^B$, then whites have much less incentive to connect to oppositional blacks (denoted by B_m), even if the latter have a lot of white friends. On the contrary, oppositional blacks want to connect to whites because of the high externalities generated from a link with a white. In particular, a white agent might not have enough incentives to build a link with a second oppositional identity black because the indirect externalities that she receives from the other white who is already connected to the black community are large enough. With fixed δ^B, this can happen when the direct externalities δ^W whites exert are very large.[20] In this case, because of mutual consent, there cannot be two oppositional identity blacks (i.e. type B_m). However, when δ^W is not too large compared to δ^B, then coordination problems are less severe and bridge links are easier to form. In this case, more than one blacks with oppositional identity can exist in equilibrium. Observe that Proposition 9 holds if the size of the communities are the same, i.e. $n^B = n^W = n/2$. Thus, this example highlights the crucial role of coordination problems and mutual consent in friendship relationships.

[20]This can be seen in condition (48) in the proof of Proposition 9.

5 Social Norms

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

$$u_i(g) = \sum_{j \in N \backslash \{i\}} b_{ij} - \sum_{j \in N_i(\mathbf{g})} c_{ij} \tag{19}$$

In this utility function, the benefits from connections are

$$b_{ij} = \max_{p_{ij} \in P_{ij}(g)} \omega(p_{ij}) \tag{20}$$

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:[21]

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

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

The interpretation of (19) 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 whites, the benefits of direct connections is δ whatever the race of the friend. For blacks, the benefits of a direct connection with a white is δ while with a black is $s(\left|Q_{ij}^W\right|)\delta$. This function, which is between 0 and 1, aims at capturing the idea of *social norms* and *social norms* from the black community. If some blacks decide to have a lot of white friends, there is a "penalty" from the black community. The function $s(\left|Q_{ij}^W\right|)$ is decreasing in the percentage of i's white friends, which means that blacks obtain less and less benefits from their direct black friends, the higher is their number of white friends. Interestingly, indirect connections are also affected by the social norms $s(.)$ for both blacks and whites. 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

[21] 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.

ridiculed him for that, they ridiculed him for going away to school, they ridiculed him for turning white. I know because he told me they did." (Anson, 1985, p. 205). In his autobiographical essay, Rodriguez (1982) told us about his own story as a Mexican-American from Sacramento who went to college and for whom English became his dominant language. His (extended) family considered him increasingly alien and as he put it: "*Pocho*, they called me. Sometimes, playfully, teasingly, using the tender diminutive −*mi pochito*. Sometimes not so playfully, mockingly, *Pocho* (Rodriguez (1982, p. 29).[22] These two stories of a black person labeled a white man by his black neighbors and an Hispanic labeled a "gringo" by his extended family are strikingly similar and illustrate the idea of social sanctions and social norms imposed by their own communities.[23]

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

$$\begin{aligned} s(\left|Q_{ij}^{W}\right|) &= 1-(1-\underline{s})\left|Q_{ij}^{W}\right| \\ &= 1-(1-\underline{s})\left|q_i^W-q_j^W\right| \end{aligned} \tag{21}$$

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

To understand the role of sanctions/norms in the utility function, let us calculate the benefits,[24]i.e. $\sum_{j\in N\setminus\{i\}} b_{ij}$ in (19), 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:

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

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

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

[24]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.

while with sanctions, there are given by:

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

Indeed, B_m receives a direct sanction from individual B_0, which is equal to $(1-\underline{s})\left(\frac{n^W}{n^W+1} - 0\right)\delta$, and an indirect sanction from the $n^B - 1$ individuals who are direct friends of B_0 equal to $(1-\underline{s})\left(\frac{n^W}{n^W+1} - 0\right)\delta^2$. The sanctions are here maximal because $\left|q_i^W - q_j^W\right|$ is the maximal value one can obtain since B_m has before the possible link with B_0 only white friends while B_0 has no white friend at all. This benefit function aims at capturing the fact that becoming totally assimilated to the white culture (i.e. having only white 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 black 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 10 *Assume (3). If the utility is defined by (19) and social norms by (21), then if*

$$C > \left[1 - \frac{(1-\underline{s})\, n^W}{n^W+1}\right]\left(\frac{n^W+1}{n^W-1}\right)\left[\delta + \left(n^B - 1\right)\delta^2\right] - \left(\frac{n^W+1}{n^W-1}\right)c \tag{22}$$

the network described in Figure 8 is pairwise stable. Condition (22) is less restrictive than (8), 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 5), "attractiveness" was crucial for the result. When social norms are introduced, having white friends increase even more the distance between the *assimilated* black B_m and the *separated* blacks 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 white community but also because she is getting less benefits when interacting with blacks. Interestingly, for the white community, B_m is still considered as a "white" person since the cost of interacting with her is still c. For the black community, B_m is less "valuable" in terms of friend than any other black.

6 Social welfare: Integration versus segregation

In this section, we consider some welfare implications of our benchmark model of Section 3. We have previously focused on how decentralized behavior 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.[25] 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 11 *Assume $c < \delta - \delta^2$ and (4). If*

$$n^B\left(n^W - 1\right)\left(n^B - 1\right) \leq (n-1)^2 \tag{23}$$

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 community active.

7 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

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

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 modelled these decisions through a precise network structure that shapes social interactions and the exposure and assimilation of others' differences.

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.[26] It would be interesting to encompass 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.[27] 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.

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.

[26] See, for example, Coleman (1961), where a taxonomy of identities that adolescents adopt in US high-schools is provided, and Akerlof and Kranton (2002), for an economic analysis of identity in schooling.

[27] See Akerlof (2008) for a model that encompasses both an identity and a social capital choice.

[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, R. (2008), "A theory of social motivation," Unpublished manuscript, Harvard University.

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

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

[9] Bala, V. and S. Goyal (2000), "A non-cooperative model of network formation," *Econometrica* 68, 1181-1229.

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

[11] Battu, H. and Y. Zenou (2009). "Oppositional identities and employment for ethnic minorities. Evidence for England," *Economic Journal*, forthcoming.

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

[15] Bramoullé, Y. and R. Kranton (2007), "Risk sharing across communities," *American Economic Review Papers and Proceedings* 97, 70-74.

[16] Bramoullé, Y. and R. Kranton (2008), "The formation of risk-sharing networks," Unpublished manuscript, Duke University.

[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] Calvó-Armengol, A. and R. Ilkiliç (2009), "Pairwise stability and Nash equilibria in network formation," *International Journal of Game Theory* 38, 51-79.

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

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

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

[23] 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.

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

[25] 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.

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

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

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

[29] Fryer, R.G. and P. Torelli (2005), "An empirical analysis of 'acting white'," NBER Working Paper No. 11334.

[30] Galeotti, A. (2006), "One-way flow networks: The role of heterogeneity," *Economic Theory* 29, 163-179.

[31] Galeotti, A., Goyal, S. and J. Kamphorst (2006), "Network formation with heterogeneous players," *Games and Economic Behavior* 54, 353-372.

[32] Golub, B. and Jackson M. (2008), "How Homophily affects Diffusion and Learning in Networks," Unpublished manuscript, Stanford University.

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

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

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

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

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

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

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

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

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

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

[43] 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.

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

[45] 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.

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

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

[48] 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.

[49] Nekby, L. and M. Rödin (2009), "Acculturation identity and employment among second and middle generation immigrants," *Journal of Economic Psychology*, forthcoming.

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

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

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

[53] Pin, P., Franz, S. and M. Marsili (2008), "Opportunity and choice in social networks," Unpublished manuscript, The Abdus Salam International Centre for Theoretical Physics.

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

[55] 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.

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

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

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

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

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

APPENDIX

Proof of Proposition 1. Condition (3) ensures that all individuals of same type want to be connected among them. Therefore both communities are fully intra-connected.

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

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

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

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

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

is the indirect benefit derived from the diminishing costs of maintaining a link with a black person, once this new link is formed. Before forming the new link, the proportion of white friends among all white person's friends is $\frac{n^W-1}{n-3}$. Once the new link is created, this proportion diminishes to $\frac{n^W-1}{n-2}$, and this implies a decrease in the cost of maintaining the link with the $n^B - 1$ black persons from $c + \left(\frac{n^B-1}{n-2}\right)\left(\frac{n^W-1}{n-3}\right)C$ to $c + \left(\frac{n^B-1}{n-2}\right)\left(\frac{n^W-1}{n-2}\right)C$. The third term in (24) accounts for this difference in costs.

The inequality (24) is equivalent to

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

which is always true since $n^B - 1 - (n-3)\left(n^W - 1\right) < 0$ and $\delta - \delta^2 - c > 0$. Similarly, because of symmetry, the condition that guarantees that there is *no* gain in utility for a black person to severe a link with a white person is

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

which is always true since $n^B - 1 - (n-3)\left(n^W - 1\right) < 0$ and $\delta - \delta^2 - c > 0$. As a result, *complete integration* between communities is always an equilibrium network.

(*ii*) Let us show that *complete segregation* between communities is an equilibrium network. There is *no* gain in utility for a white person to establish a link with a black person, who is necessarily connected to the rest of the black community, if:

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

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

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

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

(*iii*) Therefore, there is necessary *partial integration* for all networks when condition (4) does not hold. ▮

Proof of Proposition 2

(*i*) Let us denote W_1 (resp. B_1) the unique white (resp. unique black) agent involved in the bridge link. Firstly, the agent W_1 has incentives to form a link with B_1 iff

$$\delta + \left(n^B - 1\right)\delta^2 - c > C \tag{25}$$

Similarly, the agent B_1 has incentives to form the link with W_1 iff

$$\delta + \left(n^W - 1\right)\delta^2 - c > C \tag{26}$$

Since $n^W \geq n^B$, 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 (25) 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 W_1 and B_1 has incentives to form a link.

The white agent W_1 does not have incentives to form a link with a black different than B_1 iff

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

which is equivalent to:

$$\left(\frac{n^W}{n^W - 2}\right)\left(\delta - \delta^2 - c\right) < C \tag{27}$$

The black agent B_1 does not have incentives to form a link with a white different than W_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:

$$\left(\frac{n^B}{n^B-2}\right)\left(\delta-\delta^2-c\right)<C \tag{28}$$

Because of mutual consent, and since $n^W \geq n^B$ only condition (27) is required.

Any other white agent different than W_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

$$\left(\frac{n^B}{n^B-1}\right)\left[\delta+\left(n^B-2\right)\delta^2-\left(n^B-1\right)\delta^3-c\right]<C \tag{29}$$

Because of symmetry, any other black agent different than B_1 does not have incentives to form a link with W_1 iff

$$\left(\frac{n^W}{n^W-1}\right)\left[\delta+\left(n^W-2\right)\delta^2-\left(n^W-1\right)\delta^3-c\right]<C \tag{30}$$

Finally, any white agent other than W_1 does not have incentives to form a link with a black 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

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

Any black agent other than B_1 does not have incentives to form a link with a white different than W_1 iff

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

which is equivalent to

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

Because of mutual consent, only one of the conditions among (31) and (32) has to hold. Observe that if either (29) or (30) holds, then either (31) or (32) hold too.

Gathering everything together the required conditions for the network to be pairwise stable are given by (25), (27), (29) and (30).

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

$$\delta-\delta^3-c>C \tag{33}$$

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

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

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

which is equivalent to

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

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

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

Because of mutual consent, only (34) or (35) 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 black agent with a bridge link does not have incentives to build a link with a white 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

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

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

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

Because of mutual consent, only (36) or (37) is needed. since the second of these conditions is less restrictive, it suffices to ensure that this type of link is not formed.

Since

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

condition (34) implies (36).

Hence, gathering everything together we obtain that the two required conditions are (33) and (34). ∎

Proof of Proposition 4

We call oppositional black, the black agent that has a bridge link with each of the members of the white community. We denote this agent by B_m. The oppositional black individual B_m does not want to severe any of his bridge links iff

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

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

which is equivalent to

$$\frac{(n-2)(n-3)}{\left(n^B - 1\right)\left(n^B - 2\right)} \left(\delta - \delta^2 - c\right) > C \tag{38}$$

A white agent does not want to severe his bridge link with the oppositional black 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

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

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

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

which is equivalent to

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

A white agent does not have incentives to connect with a non-oppositional black B_0 iff

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

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

which is equivalent to

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

The conditions (38) and (39) have to hold. Because of mutual consent, only one of the conditions (40) and (41) is required. Condition (40) is less restrictive than the last one, and hence, the set of required conditions are (38), (39) and (40), that can hold at the same time. ∎

Proof of Proposition 5

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

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

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

which is equivalent to

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

The black individual B_m does not want to severe a link with a white individual iff

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

which is equivalent to

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

This is always true because of assumption (3).

Observe that, because of (3), none of the blacks B_0 would like to severe a link with a B_0. Because of mutual consent, when condition (8) holds, they cannot have a link with B_m because B_m does not want to.

A white will not severe a link with B_m iff

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

which always true because of (3).

Finally, because of (3), it is easy to verify that a white will never want to severe a link with another white.

Therefore, condition (8) is enough to guarantee that the network described by Figure 6 is pairwise stable. ▮

Proof of Proposition 6

A white agent does not have incentives to severe a link with a black agent whenever

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

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

On the other hand, a white agent does not have incentives to create a link with another white 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 black agent. ▮

Proof of Proposition 7

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

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

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

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

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

is the indirect benefit derived from the diminishing costs of maintaining a link with a black person, once this new link is formed. Before forming the new link, the proportion of white friends among all white person's friends is $\frac{n^W-1}{n-3}$. Once the new link is created, this proportion diminishes to $\frac{n^W-1}{n-2}$, and this implies a decrease in the cost of maintaining the link with the $n^B - 1$ black persons from $c + \left(\frac{n^B-1}{n-2}\right)\left(\frac{n^W-1}{n-3}\right) C$ to $c + \left(\frac{n^B-1}{n-2}\right)\left(\frac{n^W-1}{n-2}\right) C$. The third term in (42) accounts for this difference in costs.

The inequality (42) is equivalent to

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

The RHS of this inequality is negative iff:

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

A sufficient condition is thus:

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

Similarly, because of symmetry, the condition that guarantees that there is *no* gain in utility for a black person to severe a link with a white person is

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

which is always true if (11) is satisfied. As a result, *complete integration* between communities is always an equilibrium network if (11) is satisfied.

(*ii*) Let us show that *complete segregation* between communities is an equilibrium network. There is *no* gain in utility for a white person to establish a link with a black person, who is necessarily connected to the rest of the black community, if:

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

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

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

Since $n^W \geq n^B$, 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 black person can obtain by making a link to another black is $\delta - \delta^2$. The highest cost for a black i to have a link with another black is found by

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

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

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

So

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

This implies that the worst case scenario is

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

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

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

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

Since

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

Then the condition for both blacks and whites is

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

Putting together (4) and this condition leads to

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

■

Proof of Proposition 8

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

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

Similarly, the black center does not have incentives to build a link with the white center iff

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

Because of mutual consent, only the first one, that coincides with (14), needs to hold.

If the white center has no incentives to build a link with the periphery of the other community, this suffices to ensure that this agent does no want to form a link with anybody from the other community, and the result follows.

(*ii*) Firstly, if

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

then neither the white center nor the black center has incentives to severe the bridge link that connects them.

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

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

Observe that this is satisfied by assumption.

Similarly, the black center does not have incentives to connect with a peripheral agent of the white 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 white peripheral agent does not have incentives to connect with a black peripheral iff

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

Because of symmetry, this same condition ensures that a black peripheral agent does not have incentives to connect with a white 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 white peripheral agent does not have incentives to severe his bridge link with a black peripheral iff

$$\delta - \delta^5 + \delta^2 - \delta^4 - c > C \tag{43}$$

This same condition ensures that a black peripheral agent does not have incentives to severe his bridge link with a white peripheral agent.

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

$$\delta - \delta^3 < c \tag{44}$$

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

The white center does not have incentives to form a link with the black center iff

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

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

Observe that if (44) then (45) immediately follows.

The white center does not have incentives to form a link with a black peripheral agent iff

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

and this same condition ensures that the black center does not have incentives to form a link with a white peripheral agent. Because of (44), this last inequality trivially holds. 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 white peripheral agent does not have incentives to form a bridge link with another black peripheral agent iff

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

which is equivalent to:

$$4\left[c - \left(\delta - \delta^3\right)\right] > C \tag{46}$$

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

Hence, the required conditions are (43), (44) and (46).

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

$$\delta - \delta^3 - c > C \tag{47}$$

A white peripheral agent and a black peripheral agent don't have incentives to severe the link that connects them

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

which is equivalent to:

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

which trivially holds if (47) holds too.

A peripheral agent of one of the communities does not have incentives to form a link with the center of the other community iff

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

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

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

which is a condition that is trivially satisfied given the assumption that $c > \delta - \delta^2$. 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.

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

▮

Proof of Proposition 9

We assume that (18) holds, so that each community is always fully connected.

A white does not want to severe a link with an oppositional black B_m iff:

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

which is equivalent to:

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

A white does not want to create a link with any other black B_0 iff:

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

which is equivalent to:

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

An oppositional black B_m does not have incentives to severe any bridge link iff:

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

which is equivalent to

$$C < \frac{\left[\left(1 - \delta^W\right)\delta^W - c\right] n^W (n-2)(n-3)}{\left(n^B - 1\right)\left(n^B - 2\right)\left(n^W - 1\right)} \tag{50}$$

A non-oppositional black B_0 does not have incentives to create a link with a white iff

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

which is equivalent to

$$C > \left[\left(1 - \delta^B\right)\delta^W - c\right]\left(\frac{n^W + 1}{n^W - 1}\right) \tag{51}$$

Because of mutual consent only three conditions have to hold, that is either (48), (49), and (50) or (48), (50), and (51).

Let us start with conditions (48), (49), and (50). For these three conditions to be true, it has to be in particular that

$$\frac{\left(n^W + 1\right) n^W (n-2)\left[\left(1 - \delta^B\right)\delta^B - c\right]}{\left(n^W - 1\right)\left[n^W (n-2) - 2\left(n^B - 1\right)\right]} < \left(\frac{n-2}{n^B - 1}\right)\left(\frac{n^W}{n^W - 2}\right)\left[\delta^B\left(1 - \delta^W\right) - c\right] \tag{52}$$

which is equivalent to:

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

It is easy to see that

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

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

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

As a result, firstly, we can ensure that if $\delta^W=\delta^B$ the network analyzed is pairwise stable, and by a continuity argument this network remains pairwise stable for some range of parameters in which $\delta^W>\delta^B$. Secondly, for fixed n^W and n^B, if δ^W is very large compared to δ^B, inequality (52) might not hold.

Let on now consider conditions (48), (50), and (51). For these three conditions to be true, it has to be in particular that

$$\left[\left(1-\delta^B\right)\delta^W-c\right]\left(\frac{n^W+1}{n^W-1}\right)<\left(\frac{n-2}{n^B-1}\right)\left(\frac{n^W}{n^W-2}\right)\left[\delta^B\left(1-\delta^W\right)-c\right] \tag{53}$$

which is equivalent to

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

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

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

which ensures that, indeed, the four conditions (48), (49), (50), and (51) are compatible at the same time when $\delta^W=\delta^B$ and also, by a continuity argument, for a range of parameters with $\delta^W>\delta^B$. Anyhow, for fixed n^W and n^B, if δ^W is very large compared to δ^B, inequality (53) might neither hold. ▮

Proof of Proposition 10

Consider the network described in Figure 8. There are n^W individuals who are all connected with each other. There is one black B_m who is connected to all whites and is not connected to any other black B_0. All the other n^B-1 blacks are fully connected with each other.

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

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

which is equivalent to

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

which is (22).

As in the proof of Proposition 5, condition (3) guarantees that no individual in this network wants to create or severe a link.

Let us now show that this condition is less restrictive that (8), i.e. the one in Proposition 5 where there are no social norms. We want to show that

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

which is equivalent to

$$\underline{s} < 1$$

which is always true by definition. ▮

Proof of Proposition 11

The total surplus for complete segregation is equal to:

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

while the total surplus for complete integration is given by:

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

Segregation is better if and only if:

$$\delta \leq \left[c+\frac{\left(n^W-1\right)\left(n^B-1\right)}{(n-1)^2}C\right]\left(1-\frac{n}{2n^W n^B}\right) \tag{54}$$

It is easy to check that $1 > \frac{n}{2n^W n^B}$.

We know that for C large enough, segregation dominates integration. What happens when C is smaller? Let's take the smallest value that C can take, that is $\delta + \left(n^B - 1\right)\delta^2 - c$ and see if integration dominates segregation. The condition (54) is now given by:

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

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

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

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

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

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

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

So a sufficient condition is that

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

which is equivalent to (23). ∎