I Z A Institute of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 11436

Conformism, Social Norms and the Dynamics of Assimilation

Gonzalo Olcina
Fabrizio Panebianco
Yves Zenou

MARCH 2018

I Z A Institute of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 11436

Conformism, Social Norms and the Dynamics of Assimilation

Gonzalo Olcina
Universitat de Valencia and ERI-CES

Fabrizio Panebianco
IGIER, Bocconi University

Yves Zenou
Monash University, CEPR and IZA

MARCH 2018

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

Conformism, Social Norms and the Dynamics of Assimilation*

We consider a model where each individual (or ethnic minority) is embedded in a network of relationships and decides whether or not she wants to be assimilated to the majority norm. Each individual wants her behavior to agree with her personal ideal action or norm but also wants her behavior to be as close as possible to the average assimilation behavior of her peers. We show that there is always convergence to a steady-state and characterize it. We also show that different assimilation norms may emerge in steady state depending on the structure of the network. We then consider the role of cultural and government leaders in the assimilation process of ethnic minorities and an optimal tax/subsidy policy which aim is to reach a certain level of assimilation in the population.

JEL Classification: D83, D85, J15, Z13

Keywords: assimilation, networks, social norms, peer pressure, cultural leader

Corresponding author:
Yves Zenou
Department of Economics
Monash University
Caulfield East VIC 3145
Australia
E-mail: yves.zenou@monash.edu

* We thank Carlos Bethencourt Marrero, Benjamin Golub and Theodoros Rapanos for helpful comments.

1 Introduction

In his book, *Assimilation, American Style*, Salins (1997) argues that an implicit contract has historically defined assimilation in America. As he puts it: "Immigrants would be welcome as full members in the American family if they agreed to abide by three simple precepts: First, they had to accept English as the national language. Second, they were expected to live by what is commonly referred to as the Protestant work ethic (to be self-reliant, hardworking, and morally upright). Third, they were expected to take pride in their American identity and believe in America's liberal democratic and egalitarian principles." Though hardly exhaustive, these three criteria certainly get at what most Americans consider essential to successful assimilation.

The same issues have been discussed and debated in Europe, especially over recent decades. According to the 2016 Eurostat statistics, 20.7 million people with non-EU citizenship are residing in the European Union. Additionally, 16 million EU citizens live outside their country of origin in a different Member State. Migration movements are on the rise both within and from outside the European Union.

The key to ensuring the best possible outcomes for both the migrants and the host countries (both in the European Union and the United States) is their successful integration and assimilation into host countries.[1] However, assimilation is often fraught with tension, competition, and conflict. There is strong evidence showing that family, peers and communities shape the individual assimilation norms and, therefore, affect assimilation decisions. In particular, there may be a conflict between an individual's assimilation choice and that of her peers and between an individual's assimilation choice and that of her family and community. For example, an ethnic minority may be torn between speaking one language at home and another at work.

In this paper, we study how these conflicting choices affect the long-run assimilation behaviors of ethnic minorities and how policies and economic incentives can affect these assimilation decisions.

We consider a model where each individual (or ethnic minority) is embedded in a *directed*

[1]Different countries have different views on the integration of immigrants. Certain countries, such as France, consider it to be a successful integration policy when immigrants leave their cultural background and are "assimilated" into the new culture. Other countries, such that the United Kingdom, consider that a successful integration policy is that immigrants can keep their original culture while also accepting the new culture (or at least not rejecting it). In this paper, we will focus on the role of "assimilation" in the "integration" of ethnic minorities. However, assimilation can be defined as in Salins (1997) or in a broader way such as, for example, by the economic success of the individuals. For example, some groups such as the Chinese in the United State or in Europe, because of their economic success, can be considered as "assimilated" even if they do not interact too much with people from the majority culture.

network of relationships and where peers are defined as *outdegrees*. Each agent decides how much she wants to assimilate to the majority norm. At time t, this decision, denoted by action x_i^t for individual i, is continuous and is between 0 (no assimilation at all) and 1 (total assimilation). Each individual i wants to minimize the distance between x_i^t and the average choice of individual i's direct peers (i.e. i's social norm) but also between x_i^t and i's assimilation norm at time t, denoted by s_i^t. At time $t+1$, the latter is determined for each individual i by a convex combination of x_i^t and s_i^t. In this framework, we study the dynamics of both assimilation choice x_i^t and norm s_i^t and their steady-state values. Basically, this is a coordination game with myopic best-reply dynamics where, in each period, the agents select best responses to last-period actions.

First, we provide a micro foundation of the DeGroot model (see DeGroot (1974)) by embedding agents with a utility function that captures both their desire to conform to the average action of their peers and to be consistent with their own social norm. We show that, in order to study the dynamics of the individual norms, it suffices to look at the (row-normalized) adjacency matrix of the network even though the process of assimilation is described by a more general matrix. We also show that convergence always occurs, independently of the network structure. We characterize the steady-state individual norms and show that they depend on the initial norms (at time $t = 0$) and on the individual position in the network. In particular, we demonstrate that, if an individual belongs to a closed communication class, then her steady-state assimilation norm will be a weighted combination of the initial norms of all agents belonging to the same communication class, where the weights are determined by the eigenvector centrality of each agent. If an agent does not belong to any closed communication class, her steady-state assimilation norm will be determined by a weighted combination of the steady-state norms of the agents in the closed communication classes for which she has links or paths to.

In this framework, we can explain why individuals from the same ethnic group can choose to adopt oppositional identities, i.e. some assimilate to the majority culture while others reject it.[2] Indeed, some people may reject the majority norm because they live in closed communities that do not favor assimilation and because some of their members (for example, cultural leaders) have a strong influence on the group. In the terminology of our model, these are individuals who have a key position in the network, i.e. who have a high eigenvector centrality. On the contrary, other individuals from the same ethnic group

[2] Studies in the United States have found that African American students in poor areas may be ambivalent about learning standard English and performing well at school because this may be regarded as "acting white" and adopting mainstream identities (Fordham and Ogbu (1986); Wilson (1987); Delpit (1995); Ogbu (1997); Ainsworth-Darnell and Downey (1998); Fryer and Torelli (2010); Patacchini and Zenou (2016)).

or religion may want to be assimilated because their communities are not isolated or, if they are isolated, their social norms are in favor of assimilation. Our model shows that these two types of behaviors can arise endogeneously in the steady-equilibrium, even for ex ante identical individuals, i.e. individuals with exactly the same characteristics. The key determinant of these assimilation choices is their position in the network and the (initial) social norms of the persons they are connected to.

Second, we show that the steady-state assimilation actions of individuals are equal to their steady-state assimilation norms so that, in steady-state, for each agent, the distance between her action and that of her peers is equal to zero and the distance between her action and her own norm is also equal to zero. This implies that total welfare is maximized in the steady-state equilibrium. This is not true at any period t outside the steady state where individuals choose actions different from their social norms.

Third, when we introduce a cultural leader (such as, for example, an Imam for the Muslim community) who is "stubborn" i.e. he is not influenced by any external opinion, then the assimilation rate of all ethnic minorities in the network will converge to the cultural leader's beliefs. We then introduce a government leader (a secular institution such as the government who promotes the cultural norm of the host society) who is also "stubborn" and wants ethnic minorities to assimilate to the majority's norm. Since there is competition between the cultural and the government leader, we determine which person in the network the government leader wants to connect to in order to undermine the role of the cultural leader. We show that government leader will pick an individual who has a low degree but a high Katz-Bonacich centrality.

Fourth, we demonstrate that the speed of convergence of the individual norms depends not only on the second largest eigenvalue of the adjacency matrix but also on the taste for conformity and on the weight put on the impact of past assimilation decisions on current individual norms.

Fifth, we show that our results depend on the way peers or links in the network are defined. If, for example, peers are defined as *indegrees* and not as outdegrees, then the results change dramatically because the definition of closed communication classes are modified. The same reasoning applies if peers are defined as mutual friends. These results are particularly important for empirical applications.

Sixth, we generalize our utility function by introducing idiosyncratic economic incentives for assimilation *via* ex ante heterogeneity, which is defined as the marginal benefits of exerting action x_i^t. We propose a new way of calculating the steady-state norms and actions and show that they do not depend on initial norms but on the ex ante heterogeneity of each agent, the taste for conformity, and the position in the network.

Finally, we derive some policy implications of our model with this generalized utility function. We determine the optimal of tax/subsidiy that needs to be given to each agent in order to reach a certain degree of long-run assimilation. For example, if the objective is that all ethnic minorities much reach an assimilation level of at least 50 percent, then we are able to calculate the level of tax/subsidiy given to each agent, which depends on her marginal benefits of assimilation and her position in the network.

The rest of the paper unfolds as follows. In the next section, we relate our paper to the relevant literature. In Section 3.1, we describe the benchmark model while, in Section 3.2, we study the dynamics of individual norms. Section 3.3 is devoted to the steady-state assimilation choices and welfare while Section 3.4 considers the role of cultural versus government leaders. Section 3.5 studies the speed of convergence of individual norms whereas, in Section 3.6, we revisit our results when peers are defined as indegrees and as mutual friends. In Section 4, we consider a more general utility function where the ex ante heterogeneity of all individuals is introduced. Section 4.1 determines the steady-state equilibrium while Section 4.2 studies the policy implications of this model. In Section 5, we propose other applications of our model. Finally, Section 6 concludes. In the Appendix, we provide the proofs of the results in the main text. We have created a not-for-publication Online Appendix. In the Online Appendix A, we provide some standard results in linear algebra. In the Online Appendix B, we study the convergence results for any possible network. In the Online Appendix C, we provide the proof of equation (14). Online Appendix D gives a detailed analysis of the speed of convergence in the benchmark model while Online Appendix E explores the implications in terms of assimilation of defining peers in different ways. Finally, Online Appendix F provides the proofs of all results stated in the Online Appendix.

2 Related literature

We now relate our paper to two main literatures where network effects matter.[3]

2.1 Diffusion and learning in networks

There is an important literature on diffusion and learning in networks.[4] Our paper is more related to the repeated linear updating (DeGroot) models and, in particular, to the papers

[3]For overviews on the economics of networks, see Jackson (2008), Benhabib et al. (2011), Ioannides (2012), Jackson (2014), Jackson and Zenou (2015), Bramoullé et al. (2016) and Jackson et al. (2017).

[4]For a recent overview of this literature, see Golub and Sadler (2016).

by DeGroot (1974), DeMarzo et al. (2003), Golub and Jackson (2010, 2012). These papers propose a tractable updating model and provide conditions on the network topology under which there is convergence of opinions and characterize the steady-state solutions. There have been different extensions of the standard DeGroot model (see the overview by Golub and Sadler (2016)) and some microeconomic foundations using myopic best-reply dynamics.

With respect to this literature, we have the following contributions: (*i*) We microfound the DeGroot model and show the long run equivalence between the adjacency matrix of the network and the real matrix driving the dynamics. Our microfoundation generalizes that of Golub and Jackson (2012), who considers a model where agents minimize the distance between own action and that of their peers.[5] (*ii*) We study the welfare properties of the steady-state equilibrium and study a model where a cultutal and a government leader with conflicting objectives compete in order determine the steady-state level of assimilation of each individual in the network. (*iii*) We calculate the speed of convergence of norms and actions and show that the process can be slower or faster than that of the standard DeGroot model (Golub and Jackson (2012)) depending of some key parameters of the model. (*iv*) We show that depending on the definition of peers, the convergence of norms and actions can dramatically differ. (*v*) We introduce a more general utility function where ex ante heterogeneity (the marginal benefit of each action is specific to each individual) is explicitly modeled and propose a new methodology to determine the convergence of norms and actions of all individuals in any given network. We also study the policy implications of this model.

2.2 Assimilation of ethnic minorities: the role of networks and cultural leaders

Assimilation of minorities in a given country is an important research topic in social sciences, in general, and in economics, in particular (see e.g. Brubaker (2001), Schalk-Soekar et al. (2004), Kahanec and Zimmermann (2011), Algan et al. (2012)). The standard explanations of whether or not immigrants assimilate to the majority culture are parents' preferences for cultural traits (Bisin and Verdier (2000)), ethnic and cultural distance to the host country (Alba and Nee (1997), Bisin et al. (2008)), previous educational background (Borjas (1985)), country of origin (Beenstock et al. (2010), Borjas (1987), Chiswick and Miller (2011)), and discrimination against immigrants (Alba and Nee (1997)).

However, despite strong empirical evidence, very few papers have studied the explicit

[5] Friedkin and Johnsen (1999) propose a learning model that bears some similarities with our model but is more specific and is a particular case of our model.

impact of social networks and social norms on the assimilation outcomes of ethnic minorities and the role of cultural leaders.

2.2.1 The role of social networks in the assimilation of ethnic minorities

From an empirical viewpoint, there is a literature that looks at the impact of social networks on the assimilation choices of immigrants and ethnic minorities. To capture network effects, most economic studies have adopted ethnic concentration/enclave as the proxy for networks of immigrants in the host country (e.g. Damm (2009); Edin et al. (2003)). Other studies have used language group or language proficiency (Bertrand et al. (2000); Chiswick and Miller (2002)). The effects on assimilation are mixed. For example, Bertrand et al. (2000) and Chiswick and Miller (2002) showed that linguistic concentration negatively influenced immigrants' labor market performance in the US. In contrast, Edin et al. (2003) find that by correcting for the endogeneity of ethnic concentration, immigrants' earnings in Sweden were positively correlated with the size of ethnic concentration. Similar results were obtained by Damm (2009) for Denmark and Maani et al. (2015) for Australia.

Other papers have measured the network more directly. Gang and Zimmermann (2000) show that ethnic network size has a positive effect on educational attainment, and a clear pattern is exhibited between countries-of-origin and education even in the second generation. Using the 2000 U.S. Census, Furtado and Theodoropoulos (2010) study whether having access to native networks, as measured by marriage to a native, increases the probability of immigrant employment. They show that, indeed, marriage to a native increases immigrant employment rates. Mouw et al. (2014) use a unique binational data on the social network connecting an immigrant sending community in Guanajuato, Mexico, to two destination areas in the United States. They test for the effect of respondents' positions in cross-border networks on their migration intentions and attitudes towards the United States using data on the opinions of their peers, their participation in cross-border and local communication networks. They find evidence of network clustering consistent with peer effects.

From a theoretical viewpoint, there are few papers that analyze the role of social networks in the assimilation behaviors of ethnic minorities. As in the DeGroot model, Brueckner and Smirnov (2007) analyze the evolution of population attributes in a simple model, where an agent's attributes are equal to the average attributes value among her acquaintances. They provide some sufficient conditions on the network structure that ensure convergence to a "melting-pot" equilibrium where attributes are uniform across agents. Brueckner and Smirnov (2008) extend this model by allowing for a more general form of the rule governing the evolution of population attributes. Basically, their model is an extension of the DeGroot model where time-varying updating matrices are considered. For a similar analysis, see, in particular, Tahbaz-Salehi and Jadbabaie (2008) and Section 19.3.4 in Golub and Sadler (2016).

Buechel et al. (2015) consider a dynamic model where boundedly rational agents update opinions by averaging over their neighbors' expressed opinions, but may misrepresent their own opinion by conforming or counter-conforming with their neighbors. This is due to the fact that an agent cannot observe the true opinions of the others but only their stated opinions. They show that an agent's social influence on the long-run group opinion is increasing in network centrality and decreasing in conformity

Verdier and Zenou (2017) propose a two-stage model where the social network play an explicit role in the assimilation process. In their model, there are strategic complementarities so that more central individuals are more likely to assimilate because they obtain more utility than less central individuals.

There is also a related literature of cultural transmission where both parents and peers affect the assimilation process (see the seminal papers by Bisin and Verdier (2000, 2001)). However, very few papers have introduced an explicit network in this literature. Exceptions include Buechel et al. (2014), Panebianco (2014) and Panebianco and Verdier (2017).

Our model is quite different to these papers because we show that social norms, individual norms, and the position in the network are key determinants of the assimilation process of ethnic minorities. In particular, we show that individuals with strong assimilation norms at home may end up being not assimilated because of their "isolated" position in the network while other minorities, belonging to closed-knit networks with social norms favorable to assimilation, may end up being assimilated even though their initial assimilation norms were not in favor of assimilation. More generally, our model highlights the role of communities, ex ante heterogeneity and parental influence in the long-run assimilation of ethnic minorities.

2.2.2 The role of cultural leaders in the assimilation of ethnic minorities

There is a small but growing literature on the role of cultural leaders in the assimilation process of immigrants.[6]

Hauk and Mueller (2015) consider a model of cultural conflict where cultural leaders supply and interpret culture. The authors explain the "clash of civilizations" or "clash of cultures" between different religions and highlight the role of cultural leaders who can amplify disagreement about cultural values.

[6] For an overview, see Prummer (2018).

Carvalho and Koyama (2016) analyze religious goods by focusing on the trade off between time and money contribution to a religious good. The cultural leader (religious authority) imposes a linear tax on income-generating activity outside the community. They show that, if economic development is sufficiently low, the cultural leader chooses a strategy of cultural resistance in every period.

Prummer and Siedlarek (2017) develop an interesting model explaining the persistent differences in the cultural traits of immigrant groups with the presence of community leaders. In their model, an individual's identity is the weighted average of the host society's culture, her own past identity as well as the identity of the rest of the community, which is her network. The leaders influence the cultural traits of their community, which have an impact on the group's earnings. They determine whether a community will be more assimilated and wealthier or less assimilated and poorer. They find that cultural transmission dynamics with two opinion influencers (the host society and the group leader in their setting) result in intermediate long-run integration outcomes for the population under study.

Verdier and Zenou (2015, 2018) also study the role of cultural leaders in the assimilation process of immigrants and focus on the interaction between two leaders with opposite objectives. They show that the presence of leaders can prevent the full integration of ethnic minorities.

Compared to this literature, we model the cultural leader in a different way: he is someone who is stubborn and does not update his beliefs. Because our model is different, in the benchmark model, this leads to the fact that all individuals follow the beliefs of this cultural leader. Moreover, we study the competition between a cultural and a government leader and show to whom in a network a government leader wants to be connected to. In the more elaborated version of model where idiosyncratic heterogeneities are introduced, we show that the cultural leader has less influence over the assimilation process of the agents in the network. We also study a policy where the government leader has to decide to whom he want to connect to and what level of subsidy one must give to agents in order for them to assimilate. We believe that we are the first to study these types of policies where networks and cultural leaders play an important role in the assimilation process of immigrants.

3 The Assimilation Choice and the Dynamics of Norms

3.1 The Assimilation Choice

Consider a set N of agents with cardinality n. Agents may represent individuals belonging to some ethnic minorities, migrants or, in a simplified setting, each $i \in N$ is the representative agent of a given community. At time $t \in \mathbb{N}$, each individual $i \in N$ chooses, simultaneously with all other individuals, an action $x_i^t \in [0, 1]$. Indeed, each agent i has to decide whether or not she wants to be assimilated to the majority culture of the host country. In particular, $x_i^t \in [0, 1]$ is the assimilation effort of individual i at time t. If $x_i^t = 0$, then individual i chooses not to be assimilated at all (i.e. chooses to be oppositional) while, if $x_i^t = 1$, she chooses to be totally assimilated to the majority culture. Clearly, the higher is x_i^t, the higher is the assimilation choice. As stated in the Introduction (see, in particular, footnote 2), there is an important literature that studies the concept of *oppositional cultures* among ethnic minorities. In this literature, ethnic groups may "choose" to adopt what are termed "oppositional" identities, that is, some actively reject the dominant ethnic (e.g., white) behavioral norms (they are oppositional, which means $x_i^t = 0$) while others assimilate to it (i.e. $x_i^t = 1$).

Each agent i at time 0 is born with an *ideal action*, or *norm*, $s_i^0 \in [0, 1]$,[7] which captures both her own type and the influence of her family in terms of cultural and ethnic values (original language, customs, etc.). This norm then evolves over time through a process that we describe below. Each $i \in N$ is then exposed to a group of peers who make different assimilation decisions. More precisely, when each individual i makes a decision about x_i^t, her behavior is driven by two competing motives. First, she wants her behavior to agree with her personal ideal action s_i^t at time t, which means that there is a *consistency* between her own norm s_i^t and her behavior x_i^t. Second, she also wants her behavior to be as close as possible to the average assimilation behavior of her peers, which implies that she is *conformist*.

Each individual is embedded in a social network $\bar{\mathbf{g}}$, which is characterized by an adjacency matrix $\bar{\mathbf{G}} \in \mathbb{R}^{N \times N}$. Consider $\mathbf{G}$ to be the row-normalization of $\bar{\mathbf{G}}$, where $\mathbf{g}$ denotes the corresponding network, so that the sum of each row in $\mathbf{G}$ is equal to 1 and $g_{ij} > 0$ if and only if i assigns a positive weight to j. In other words, we consider a *directed* network where links are *outdegrees*. There are no self-loops so that $g_{ii} = 0$. Because the network is *directed*, we do not impose any symmetry on the links in the network so that we allow for $g_{ij} \neq g_{ji}$. This simply means that any two agents can give different weights to each other. This may

[7] Our framework can be extended without any change to $s_i^t \in X \subseteq \mathbb{R}$ with X being a connected interval, and $x_i^t \in X$. Notice that actual actions and ideal actions have the same support.

simply derive from the fact that the two agents have different degrees in $\bar{\mathbf{g}}$.[8] Moreover, we allow for $g_{ij} > 0$ and $g_{ji} = 0$. For the rest of the paper, we assume $\mathbf{G}$ to be diagonalizable. We call N_i the neighborhood of i, that is $N_i := \{j \in N \mid g_{ij} > 0\}$. The average action taken by i's neighbors or peers at time t is thus given by $\sum_{j \in N_i} g_{ij} x_j^t$.

The timing of the model is as follows:

1. At the beginning of each period t, each $i \in N$ is endowed with some ideal norm s_i^t.

2. Agents choose their assimilation efforts $(x_i^t)_{i \in N}$.

3. At the end of period t, each individual i updates her ideal action.

4. The process starts again.

We model the two forces described above by assuming that each i at time t chooses x_i^t that maximizes the following utility function:[9]

$$u_i^t\left(x_i^t, s_i^t, \mathbf{G}\right) = \underbrace{-\omega\left(x_i^t - \sum_{j \in N_i} g_{ij} x_j^t\right)^2}_{\text{Conformism}} - \underbrace{\left(x_i^t - s_i^t\right)^2}_{\text{Consistency}} \qquad (1)$$

The first term on the right-hand side of (1) represents the social interaction part. It is such that each individual i wants to minimize the distance between her assimilation effort x_i^t and the average assimilation effort of her peers $\sum_{j \in N_i} g_{ij} x_j^t$. Indeed, the individual loses utility $-\omega\left(x_i^t - \sum_{j \in N_i} g_{ij} x_j^t\right)^2$ from failing to conform to the average effort of her peers, where ω is her taste for conformity. This is the standard way economists have been modeling conformity behaviors in social networks (Patacchini and Zenou (2012); Liu et al. (2014); Boucher and Fortin (2016); Boucher (2016)). The second term represents the willingness to be consistent with own ideal action. In other words, each agent cares about identity-driven ideal.[10] Specifically, each agent i derives a personal utility, which is a decreasing function of the distance between her chosen action x_i^t and her ideal action s_i^t. Observe that each agent

[8] We introduce here the (original) network $\bar{\mathbf{g}}$ and its adjacency matrix $\bar{\mathbf{G}}$ because they will be crucial in Section 3.6 below.

[9] In Section 4 below, we extend this utility function to introduce ex ante heterogeneity of each agent i.

[10] That individuals derive utility from following their personal ideals has been recognized in other analyzes of social interactions. See, for example, Akerlof (1997) and Kuran and Sandholm (2008). Even if these papers bear some similarities with our model, the main difference is that the social network is not explicitly modeled.

i puts a weight ω on her social payoff while the weight she places on her personal payoff is normalized to unity.

More generally, this utility function captures the fact that, for example, a child of an immigrant makes choices responsive to those of the host society and of peers but her behaviors may conflict with those of her differently socialized parents. In other words, there will be a tension between personal preferences and coordination with peers. For example, an ethnic minority may be torn between speaking one language at home and another at work. Solving for the first-order condition, we easily obtain:

$$x_i^t = \left(\frac{1}{1+\omega}\right) s_i^t + \left(\frac{\omega}{1+\omega}\right) \sum_{j \in N_i} g_{ij} x_j^t \tag{2}$$

In matrix form, we have:

$$\mathbf{x}^t = \left(\frac{1}{1+\omega}\right) \mathbf{s}^t + \left(\frac{\omega}{1+\omega}\right) \mathbf{G}\mathbf{x}^t$$

Setting $\theta := \omega/\left(1+\omega\right)$ and solving this equation, we obtain:

$$\mathbf{x}^t = \left(\frac{1}{1+\omega}\right) \left[\mathbf{I} - \theta\, \mathbf{G}\right]^{-1} \mathbf{s}^t =: \mathbf{b}_{\mathbf{s}^t}(g, \theta) \tag{3}$$

where $\mathbf{b}_{\mathbf{s}^t}(g, \theta)$ is the weighted Katz-Bonacich centrality (Bonacich (1987); Katz (1953); Ballester et al. (2006)). Since the largest eigenvalue of $\mathbf{G}$ is 1, then $[\mathbf{I} - \theta\, \mathbf{G}]$ is invertible and with non negative entries if and only if $\theta \equiv \omega/\left(1+\omega\right) < 1$, which is always true. There is thus a unique and interior solution to (3).

Equation (3) shows that x_i^t, the assimilation effort of each ethnic minority i at time t depends on ω, the taste for conformity, her ideal action or norm at time t, s_i^t, and her position in the network, as captured by her Katz-Bonacich centrality.

3.2 The dynamics of individual norms

We now study the dynamics of social norms and examine how it is affected by individuals' assimilation choices and the structure of the network. We make two important assumptions here. First, the network is fixed and does not change over time. What changes over time is x_i^t, the assimilation effort of each individual i, and her social norm s_i^t. Second, all individuals are myopic so that they decide upon x_i^t by only considering the instantaneous utility at time t as described in (1). This is because, as in the DeGroot literature cited in Section 2.1, we want to have a tractable model. Interestingly, there are some empirical papers (field and lab experiments) that have tested whether agents behave as in our model (DeGroot model) or in a more Bayesian way. Corazzini et al. (2012), Mueller-Frank and Neri (2013), Chandrasekhar et al. (2015), show that individuals tend to behave as in DeGroot model so that agents tend to be myopic and have limited cognition.

At time t, once individual i has chosen assimilation effort x_i^t, she can reconsider her own ideal action s_i^t and update it depending on the previous action profile. We have the following dynamic equation:

$$\mathbf{s}^{t+1} = \gamma \underbrace{\mathbf{x}^t}_{\text{Consistency}} + (1-\gamma) \underbrace{\mathbf{s}^t}_{\text{Anchoring}} \tag{4}$$

where $\gamma \in [0,1]$. The parameter γ measures the level of consistency of all agents. This source for preference change has a robust psychological foundation and has been widely used in economics (see for instance, Akerlof and Dickens (1982); Kuran and Sandholm (2008)). The dynamic equation (4) states that the evolution of social norms is a linear convex combination of the past assimilation choice and the past social norm. Indeed, the first term represents how much each individual is consistent with her own assimilation choice while the second term indicates how much she is anchored to her past norm.[11]

Define $\mathbf{M}(\theta, \mathbf{G}) := [\mathbf{I} - \theta\mathbf{G}]^{-1}$. Then, by substituting (3) into (4), we obtain:

$$\mathbf{s}^{t+1} = \left[\frac{\gamma}{(1+\omega)}\mathbf{M}(\theta, \mathbf{G}) + (1-\gamma)\mathbf{I}\right]\mathbf{s}^t \tag{5}$$

Define

$$\mathbf{T} := \frac{\gamma}{(1+\omega)}\mathbf{M}(\theta, \mathbf{G}) + (1-\gamma)\mathbf{I} \tag{6}$$

Then, the dynamic equation (5) is a time-homogenous Markov process where

$$\mathbf{s}^{t+1} = \mathbf{T}\,\mathbf{s}^t \tag{7}$$

This transition mechanism takes into account each agent's ideal action and own equilibrium actions. The latter, as shown by (3), are depending on the network structure and on the position of each agent in the network (captured by their Katz-Bonacich centrality). The *limiting beliefs* can be calculated as a function of the initial beliefs and weights. They are

[11] Observe that this is related to the literature on self-signaling (Bem (1972), Benabou and Tirole (2004, 2006)), which assumes that agents, by observing their own behavior, progressively discover their own true norms and update them in the direction of past behavior, so that norms and behavior progressively converge because of the behavior-to-norm force.

given by:

$$\mathbf{s}^{\infty} = \lim_{t \to \infty} \mathbf{T}^t \, \mathbf{s}^0 \tag{8}$$

where $\mathbf{T}^t$ is the matrix of cumulative influences in period t.

Definition 1 *A matrix* $\mathbf{T}$ *is convergent if* $\lim_{t \to \infty} \mathbf{T}^t \, \mathbf{s}$ *exists for all vectors* $\mathbf{s} \in [0,1]^n$.

This definition of convergence requires that own norms converge for all initial vectors. Indeed, if convergence fails for some initial vector, then there will be oscillations or cycles in the updating of own norms and convergence will fail. It should be clear that convergence depends on the characteristics of matrix $\mathbf{T}$, which is a non-trivial transformation of the network adjacency matrix $\mathbf{G}$. $\mathbf{T}$ contains information on the equilibrium actions, the updating of the norms and the preference parameters of all the agents interacting in the network.

We now provide conditions for the convergence of $\mathbf{T}$ that only depend on the topological characteristics of $\mathbf{G}$, *independently* of the preference parameters. This is important since $\mathbf{G}$ is a network that can be observed while $\mathbf{T}$ is not. First, notice that $\frac{1}{(1+\omega)} \mathbf{M}(\theta, \mathbf{G})$ is row-normalized.[12] It follows that also $\mathbf{T}$ is row normalized, being a convex linear combination of two row-normalized matrices. For any $\epsilon \in (0,1)$, define

$$\mathbf{G}_{\epsilon} := \epsilon \mathbf{I} + (1 - \epsilon)\mathbf{G}. \tag{9}$$

$\mathbf{G}_{\epsilon}$ is the matrix of social interactions in which we consider weights *as if* all agents actually put some (homogenous) weight on themselves. Then we can provide the following steady state characterization.[13]

Proposition **1**

(*i*) *Assume that* $\mathbf{G}$ *is an aperiodic matrix. Then,*

$$\mathbf{s}^{\infty} = \lim_{t \to \infty} \mathbf{T}^t \, \mathbf{s}^{(0)} = \lim_{t \to \infty} \mathbf{G}^t \, \mathbf{s}^{(0)} \tag{10}$$

(*ii*) *Assume that* $\mathbf{G}$ *is a periodic matrix. Then,*

$$\mathbf{s}^{\infty} = \lim_{t \to \infty} \mathbf{T}^t \, \mathbf{s}^{(0)} = \lim_{t \to \infty} \mathbf{G}_{\epsilon}^t \, \mathbf{s}^{(0)} \tag{11}$$

[12] Notice that, since $\mathbf{G}$ is row-normalized, then $\sum_{t=0}^{\infty} \theta^t \mathbf{G}^t \cdot \mathbf{1}$ is a vector with all entries equal to $\frac{1}{1-\theta}$. Since $\theta = \frac{\omega}{1+\omega}$, then $\frac{1}{1-\theta} = 1 + \omega$. Then the sum of the entries of each row of the matrix $\mathbf{M}(\theta, \mathbf{G})$ is $1 + w$. It is then immediate to see that $\frac{1}{(1+\omega)} \mathbf{M}(\theta, \mathbf{G})$ is row-normalized.

[13] In the Online Appendix A, we provide some standard linear algebra results about irreducibility and aperiodicity of matrices.

This is an important result, which shows that in order to analyze the steady-state vector of own norms defined by the matrix $\mathbf{T}$, it is enough to look at the adjacency matrix $\mathbf{G}$ of the network. Moreover, convergence is ensured independently of whether $\mathbf{G}$ is periodic or not,[14] and independently of preference parameters ω and γ. This result is driven by the fact that the anchoring element of the norms dynamics makes each norm depends on its past value, thus breaking any possible cycle. This enables us to consider any network without restrictions so that our model always ensures convergence of social norms. We also provide an easy way to determine the steady state of the dynamics even when $\mathbf{G}$ shows cycles. Indeed, Proposition 1(ii) characterizes the steady state by considering a small perturbation of the original network.

Observe that, at any finite time t, $\mathbf{T}^t \neq \mathbf{G}^t$ and they only converge when $t \to \infty$. The result about the common asymptotic properties of matrices $\mathbf{T}$ and $\mathbf{G}$ given in Proposition 1 is based on the fact that these two matrices commute, i.e. $\mathbf{TG} = \mathbf{GT}$ (see the proof of Proposition 1 in the Appendix). Indeed, when $\mathbf{T}$ and $\mathbf{G}$ commute (which is based on the fact that $\mathbf{G}$ and $\mathbf{M}$ commute and are diagonalizable), then they have the same eigenvector $\mathbf{e}$ associated with the maximum eigenvalue, which is 1 here.[15] This implies that: $\mathbf{e}^{\mathrm{T}}\,\mathbf{T} = \mathbf{e}^{\mathrm{T}}\,\mathbf{G} = \mathbf{1}\,\mathbf{e}^{\mathrm{T}}$, which proves the convergence result. In Section 4, we show that these convergence results are also true for a generalized version of the utility (1) where ex ante heterogeneity in terms of observable characteristics or economic incentives to assimilation is added to the utility function.

3.2.1 Example

Let us provide a simple example that shows how the asymptotic properties of $\mathbf{T}$ and $\mathbf{G}$ are the same but, at the same time, how the two matrices differ during the convergence process.

Consider the (directed) network in Figure 1.

[14] With some abuse of notation, we say that $\mathbf{G}$ is aperiodic if the submatrix associated to each closed communication class is aperiodic. Closed communication classes are defined in the Online Appendix B.

[15] Since both $\mathbf{T}$ and $\mathbf{G}$ are row-normalized they both have the same largest eigenvalue equal to 1.

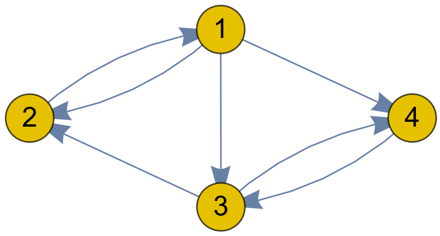

Figure 1: Network with 4 agents

It is easily verified that the (row-normalized) adjacency matrix (of outdegrees) is an irreducible and aperiodic matrix. For simplicity, set $\omega = \gamma = 0.5$. Then the following holds:

$$\lim_{t\to\infty} \mathbf{T}^t = \lim_{t\to\infty} \mathbf{G}^t = \begin{bmatrix} 0.231 & 0.231 & 0.308 & 0.231 \\ 0.231 & 0.231 & 0.308 & 0.231 \\ 0.231 & 0.231 & 0.308 & 0.231 \\ 0.231 & 0.231 & 0.308 & 0.231 \end{bmatrix}$$

While the two matrices converge to the same limit, they drastically differ at any finite time period. To show this, let us focus on $g_{12}^{[t]}$ and $t_{12}^{[t]}$ and consider Figure 2.

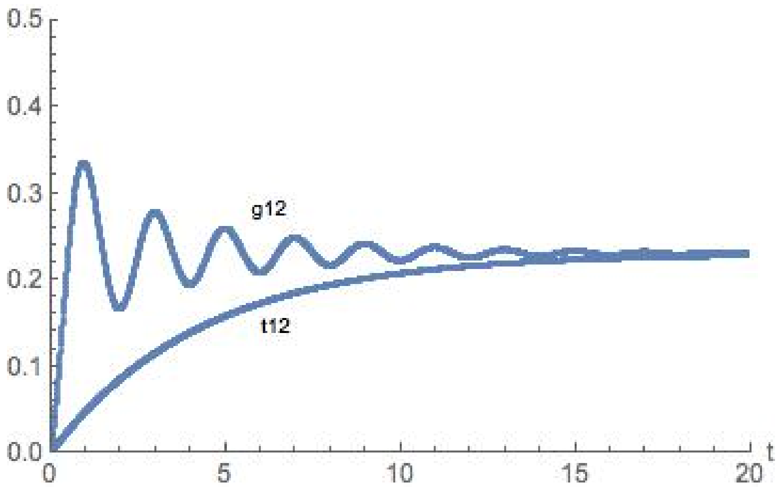

Figure 2: Convergence of $g_{12}^{[t]}$ and $t_{12}^{[t]}$

On can see that the convergence process is non monotonic for $g_{12}^{[t]}$ while, it is monotonic for $t_{12}^{[t]}$. However, they converge to the same limit in the long run.

3.2.2 The Characterization of Long Run Norms

Along the paper we will illustrate our results with reference to three different networks displayed in Figure 3, which we label network 1 (left panel), network 2 (right panel) and network 3 (bottom panel).

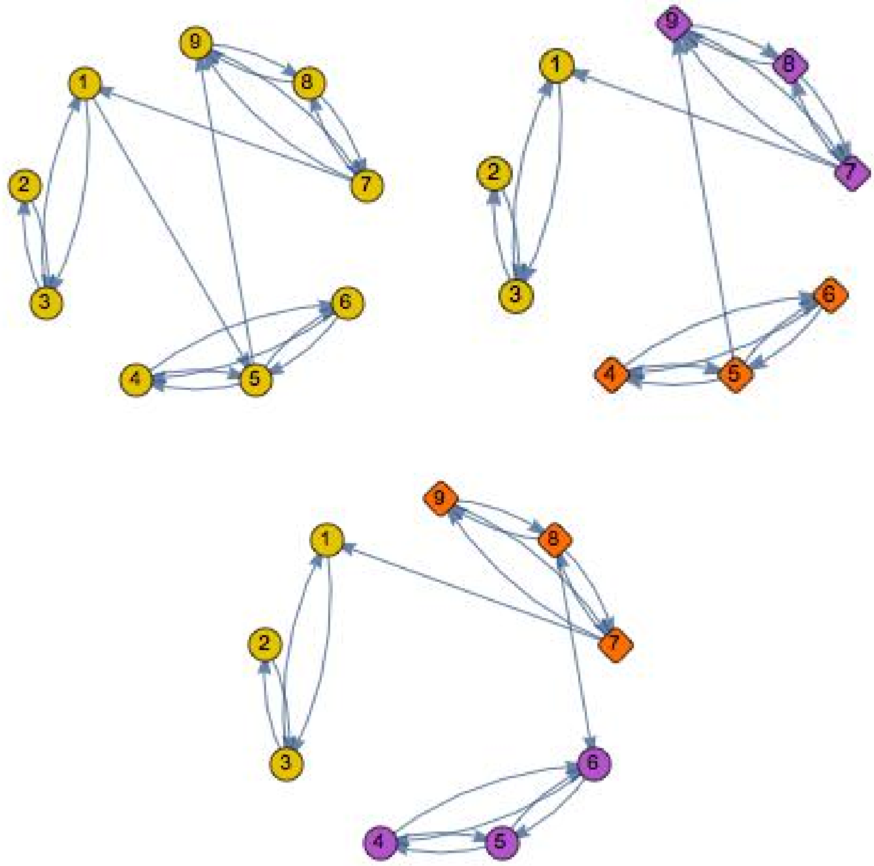

Figure 3: Three different networks

In all these networks, there are three groups of agents. These networks differ in their inter-group links. It is easily verified that $\mathbf{G_2}$ (the adjacency matrix of network 2) can be derived from $\mathbf{G_1}$ (network 1) by removing the link g_{15} in $\mathbf{G_1}$, while $\mathbf{G_3}$ (network 3) can derived from $\mathbf{G_2}$ by removing the link g_{59} and adding the link g_{86}. Interestingly, network 1 is strongly connected (i.e. all individuals belong to the same communication class), network 2 has one closed communication class where some individuals belong to it and some do not and network 3 has two closed communication classes where some individuals belong to one of them and some do not belong to any communication class (see the Online Appendix B.1 for basic definitions of communication classes).

To provide a precise characterization of steady state norms for the dynamics described in our model, in the Online Appendix B.1, we provide some important definitions on who the "influenced agents" are in a network, the concepts of communication classes and closed communication classes. In Appendices B.2, B.3, and B.4 we provide all the theoretical results on convergence for strongly connected networks, networks with one closed communication class, and any network. We report these results in the Online Appendix since they strongly rely on already existing literature. Notice, however, that with respect to the existing literature we obtain results both when the adjacency matrix $\mathbf{G}$ is aperiodic and periodic. Our results provide a precise characterization of the long run norms given any network without restriction of any sort.

In a nutshell, our results in the Online Appendix show that, for any network $\mathbf{G}$, if $\mathbf{G}$ is *strongly connected*, then long-run norms converge to a common value. If there is just *one closed communication class*, the steady-state social norms of all agents is equal to the convergence value of the social norm of this closed communication class. If there are *multiple closed communication classes*, agents who do not belong to any of these communication classes will have norms converging to a convex combination of the social norms of these closed communication classes. We provide in the Online Appendix B.2 a precise characterization for these convergence values.

In this paper, we focus on the role of the network. Without the network, similar agents randomly interacting with other agents would end up with similar long-run norms and actions. The presence of a network, instead, makes similar agents ending up with long-run norms and actions that may be very different depending on their position in the network.

3.3 Steady-state assimilation choices and welfare

Given our previous results on norm convergence, we are now able to characterize the steady-state assimilation effort choices of all individuals belonging to a given network $\mathbf{G}$. Recall first that

$$\mathbf{x}^t = \left(\frac{1}{1+\omega}\right)\left[\mathbf{I} - \theta\,\mathbf{G}\right]^{-1}\mathbf{s}^t = \left(\frac{1}{1+\omega}\right)\left[\mathbf{I} - \theta\,\mathbf{G}\right]^{-1}\mathbf{T}^t\mathbf{s}^{(\mathbf{0})} \tag{12}$$

It should be clear that the total welfare is maximized when the utility of each agent is equal to zero because, in that case, there are no losses. This is when the social norm and thus the assimilation effort of each individual is the same as those of her neighbors in the network.

Proposition 2 *In steady state, the equilibrium assimilation efforts are given by:*

$$\mathbf{x}^{\infty} = \left(\frac{1}{1+\omega}\right)[\mathbf{I} - \theta\,\mathbf{G}]^{-1} \lim_{t\to\infty} \mathbf{G^t s^0} = \mathbf{G}\mathbf{x}^{\infty} \tag{13}$$

and $\mathbf{x}^{\infty} = \mathbf{s}^{\infty}$. *Moreover, for any network* $\mathbf{G}$*, the total welfare is always maximized. If* $\mathbf{G}$ *is periodic, the same result holds by replacing* $\mathbf{G}$ *with* $\mathbf{G}_{\epsilon}$.

This proposition gives the equilibrium assimilation efforts in steady state. Consider first equation (13). The first two terms relate steady-state equilibrium actions to steady-state norms. This is derived from the limit (when t goes to infinity) of (12). This is, however, hardly informative about how steady-state actions relate to the network. The last term of (13) provides this characterization. Indeed, the steady-state vector is such that each agent chooses an action, which is the average of her own neighbors' actions.

Since actions are such that each individual conforms to the average assimilation effort of her peers, in steady state all individuals obtain a utility of zero, which clearly maximizes total welfare. The level for which the social norms converge does not have any impact on the welfare. This is because our utility only depends on conformity and on the adherence to own norm, while exerting a specific action does not have any other consequence.

Observe that the results stated in Proposition 2 just refer to steady-state actions. Let us calculate the assimiliation effort of each individual i at any moment in time t. It can be shown that:[16]

$$x_i^t - s_i^t = \sum_{j\neq i} m_{ij}(s_j^t - s_i^t) \tag{14}$$

where $\mathbf{M} = [\mathbf{I} - \theta\mathbf{G}]^{-1}$ and m_{ij} is its (i,j) entry. This means that, at time t, the difference between the assimilation effort of agent i and her assimilation norm, which can be seen as a measure of *cognitive dissonance*, is equal to the difference between the norm of her path-connected peers and her own norm. This implies that the agents who exert an effort different from their social norms are more likely to be the ones who are more connected in the network to agents who have different norms than theirs. Thus, Proposition 2 shows that, for any network $\mathbf{G}$, only in steady-state the assimilation efforts are efficient and equal to the assimilation actions since, in general, along the dynamic convergence process, welfare is not maximal.

Assume now that the planner is able to shape norms costlessly. Which vector of norms should she choose in order to maximize total welfare at any moment in time? The following proposition gives a clear answer to this question:

[16]See Online Appendix C for the derivation of (14).

Proposition 3 *Denote* $\mathbf{A} \equiv \frac{1}{1+\omega}\mathbf{G}[\mathbf{I}-\theta\mathbf{G}]^{-1}$. *At any time* t, *by setting a norm* $\mathbf{s}^{*t} = \mathbf{e}(\mathbf{A})$ *to each individual, the planner maximizes total welfare and the equilibrium utility is equal to zero.*

This is an interesting result, which shows that, if we add a first stage where the planner decides the level of the norm for each agent at each period t, then each agent will choose an action $\mathbf{x}^t$ that maximizes total welfare. This means that the dynamics of $\mathbf{s}^t$ is not anymore governed by equation (4) but by the value $\mathbf{s}^{*t}$ determined by the planner at each period of time t.

Example 1 *To illustrate these results, consider network 3 in Figure 3. Appendix B.4.1 shows that, when the initial norms are given by*

$$[\mathbf{s}^{(\mathbf{0})}]^T = \begin{bmatrix} 0.1 & 0.2 & 0.3 & 0.4 & 0.5 & 0.6 & 0.7 & 0.8 & 0.9 \end{bmatrix} \tag{15}$$

then the steady-state norms are equal to

$$[\mathbf{s}^{(\infty)}]^T = \begin{bmatrix} 0.225 & 0.225 & 0.225 & 0.5 & 0.5 & 0.5 & 0.328 & 0.397 & 0.363 \end{bmatrix} \tag{16}$$

In this example, agents 1, 2 and 3 belong to the first closed communication class C_1, *agents 7, 8 and 9 to the second closed communication class* C_2 *while agents 4, 5, and 6 do not belong to any closed communication class. The steady-state assimilation norms and actions are determined by the position of each agent in the network and whether they belong or not to a closed communication class. Let us now study the whole dynamics of the assimilation efforts of individuals 2 (who belongs to* C_1*), 4 (who belongs to* C_2*), and 9 (who does not belong to any closed communication class). Figure 4 reports their dynamics, which can be different to that of the assimilation social norms since* $\mathbf{x}^t = \mathbf{s}^t$ *in only true in steady state, i.e. when* t *tends to* ∞.

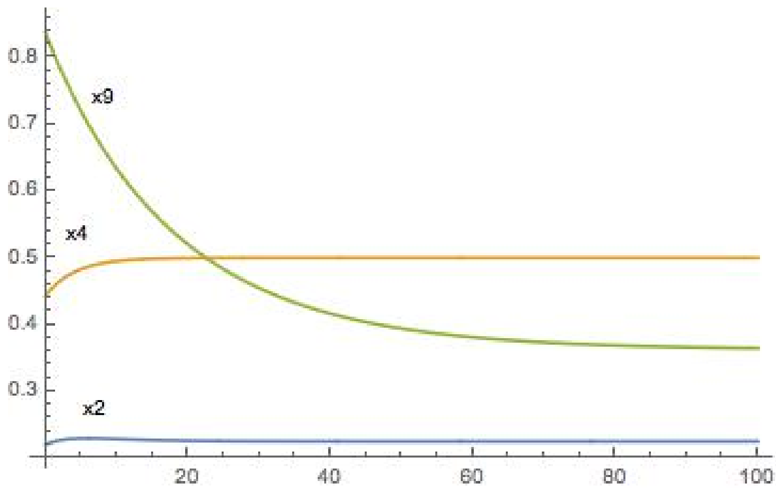

Figure 4: Convergence of actions with two communication classes

We can see, however, that the dynamics of the assimilation efforts follow closely the dynamics of the social norms (see Figure B.3 in Appendix B.4). What is interesting here is that the assimilation efforts of individuals 2, 4 and 9, who belong to different groups, have different dynamics and end up with different values in steady state. Indeed, individual 2, who belongs to the first closed communication class C_1*, will converge to the average assimilation effort of* C_1*. Similarly, individual 4, who belongs to the second closed communication class* C_2*, will converge to the average assimilation effort of* C_2*. Finally, for individual 9, who does not belong to any closed communication class, her assimilation effort will converge to some convex combination of that of the two communcation classes* C_1 *and* C_2*. More interestingly, individual 9, who starts with a very high social norm (0.9), ends up having a steady-state assimilation effort that is below that of individual 4, who starts with a much lower social norm. This is an important result in terms of assimilation choices. Individual 9, who is peripheral in the network and who inherited a high assimilation norm from her parents (who were assimilated themselves), end up choosing to be less assimilated (*$x_9^\infty = 0.363$*) because of her position in the network and the influence of her peers with whom she wants to be as close as possible in terms of assimilation choices.*

3.4 Policy implications

In this section, we study how the role of community cultural leaders influence the assimilation choices of ethnic minorities and how a policymaker can undermine the (negative) influence in terms of assimilation of the community leader by targeting specific agents in the network.

3.4.1 Cultural leaders

Let us now study the role of cultural leaders in the assimilation process of ethnic minorities. Consider a "stubborn" cultural leader CL (such as an Imam for the Muslim community), which is not influenced by external opinion and therefore does not update his beliefs. The social norm (and action) in terms of assimilation of the cultural leader is exogenous and equal to $s_{CL} = x_{CL} < 1$. This does not change over time so that $s_{CL}^0 = s_{CL}^t = s_{CL}^\infty$, $\forall t$. If, for example, his initial norm is $s_{CL}^0 = 0.1$ (not favorable to assimilation), then he will not change his opinion in the long run. To model this, we assume that the cultural leader is only linked to himself and nobody else. At the same time, some or potentially all agents in the network are linked to him.

If the original network $\mathbf{G}$ is *strongly connected*, the presence of a cultural leader creates a new network with just one closed communication class, which will only consist of the "stubborn" cultural leader CL. It is then straightforward to see that all agents in the network will converge to the social norm s_{CL}^0 of the cultural leader.

Assume now that there are *several closed communication classes* in the original network $\mathbf{G}$. These different communication classes may represent different communities. If the cultural leader is linked to at least one agent in each of these closed communication classes, then, as before, all social norms will converge to that of the cultural leader. If this is not the case, then the social norm of each agent in the closed communication class in which no one is linked to the cultural leader will just converge to a weighted average of the initial norms of the agents in this closed communication class. For all the other closed communication classes where at least an agent is linked to the cultural leader, they will adopt the cultural leader's norm. If we now consider all agents who do *not* belong to any closed communication class, then, even if these agents are not connected to the cultural leader, their steady-state norms will be a convex combination between the (stubborn) cultural leader's norm and the norm of agents in the other closed communication classes.

3.4.2 Cultural leaders versus government leaders

Let us now consider a policy where a planner (or a government) wants to thwart the influence of a cultural leader on the assimilation process of ethnic minorities. To achieve this, the planner wants to introduce a "stubborn" government leader that promotes the cultural norm of the host society (think of a secular institution in the religious example) in the network. As for the cultural leader, the government leader does not have any link with any other agent in the network but himself while other agents may be connected to him.

The government must decide to which person in the network this government leader should be linked to (link directed from the targeted agent to the leader) in order to maximize aggregate assimilation in the network. Observe that the best policy will be to link the government leader to the cultural leader but clearly this may not be a feasible policy since these two leaders are competing with each other because they have conflicting objectives. Observe also that if the planner had no budget constraint, then, apart from the cultural leader, she will connect the government leader to all agents in the network. We here assume that there is some budget constraint and that the planner needs to target only one agent in the network. To whom the government leader should be connected in order to maximize the assimilation process of the ethnic minorities?

The timing is as follows. In the first stage, the cultural leader, whose exogenous norm is $s_{CL}^0 = s_{CL}^t = s_{CL}^\infty$, $\forall t$, is already located in the network with links arbitrarily chosen. In the second stage, the government leader whose exogenous norm is $s_{GL}^0 = s_{GL}^t = s_{GL}^\infty$, $\forall t$, has to decide to which agent (apart from the cultural leader) in the network he wants to connect to. We assume that $s_{CL} < s_{GL}$.

Denote by $i = GL$ the government leader, by $j = CL$ the cultural leader and by k the agent to whom the planner wants the government leader to be connected to. Denote by $\mathbf{G}_{CL}$ the network $\mathbf{G}$ augmented by the presence of the cultural leader. Also, denote by $\mathbf{G}_{CL} + ik$ the network $\mathbf{G}_{CL}$ where the link ik between the government leader i and agent k has been added, and by d_k the (out)degree of agent k in the network $\mathbf{G}_{CL} + ik$. Denote by $\mathbf{Q}_{CL} + ik$ the matrix of weights of $\mathbf{G}_{CL} + ik$ when just the rows and columns of original agents (neither government nor cultural leader) are considered, and by $m_{ij}(\mathbf{Q}_{CL} + ik)$ the (i, j) cell of the matrix: $\mathbf{M}(\mathbf{Q}_{CL} + ik) = \sum_{p=0}^{\infty} (\mathbf{Q}_{CL} + ik)^p = [\mathbf{I} - (\mathbf{Q}_{CL} + ik)]^{-1}$. Observe that $\mathbf{M}(\mathbf{Q}_{CL} + ik)$ corresponds to the Katz-Bonacich centrality (Katz (1953), Bonacich (1987)) of the network $\mathbf{Q}_{CL} + ik$.

Proposition 4 *When the network $\mathbf{G}_{CL}$ is such that the "stubborn" cultural leader $j = CL$ is the unique closed communication class, in order to maximize aggregate assimilation in the network, the planner wants her own "stubborn" government leader $i = GL$ to target the agent k that maximizes:*

$$\frac{1}{d_k} \sum_{l=1}^{n} m_{lk} \left(\mathbf{Q}_{CL} + ik \right) \tag{17}$$

This proposition characterizes the optimal targeting choice of the government leader that undermines the influence of the cultural leader in the assimilation process of all agents in the network since, without intervention, all agents will converge to the social norm of the cultural leader.

This proposition shows that the policymaker would like to link his own leader to the agent k with the *lowest degree* but with the *highest Katz-Bonacich in-centrality.* The reason is that the agent with the highest Katz-Bonacich in-centrality is the one who has the largest cascade impact on the overall network. Indeed, this in-centrality counts how many walks an agent has starting from herself and spreading to every other agent in the network. However, since the link ik is just among the several links k has, the policymaker wants to avoid that the weight of the link created is low. This can be done by choosing the agent with a low degree. In other words, the policymaker would like to link the government leader to an agent whose norms are directly or indirectly imitated by many others, but who himself imitates very few agents. Let us illustrate this result with an example.

EXAMPLE **2** *Consider the network in Figure 1. Assume that, together with the four existing agents, there is also a cultural leader who has a social norm equal to:* $s_{CL}^0 = s_{CL}^t = s_{CL}^\infty = 0$, $\forall t$, *i.e. zero assimilation. We assume that all four agents in the network are connected to the cultural leader but the latter is only connected to himself. Then, without a government leader, all four agents in the network will end up with zero assimilation since the cultural leader is the only closed communication class (the original network is strongly connected network). We now need to decide to whom the government leader should be linked to when his objective is to maximize aggregate assimilation in the network. Assume that the social norm of the government leader is equal to:* $s_{GL}^0 = s_{GL}^t = s_{GL}^\infty = 1$, $\forall t$, *i.e. total assimilation.*

The adjacency matrices of the original network $\mathbf{G}$ *(Figure 1) and of the network with the cultural leader only* $\mathbf{G}_{CL}$ *are given by:*

$$\mathbf{G} = \begin{bmatrix} 0 & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ 1 & 0 & 0 & 0 \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 0 & 1 & 0 \end{bmatrix} ; \ \mathbf{G}_{CL} = \begin{bmatrix} 0 & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{2} & 0 & 0 & 0 & \frac{1}{2} \\ 0 & \frac{1}{3} & 0 & \frac{1}{3} & \frac{1}{3} \\ 0 & 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

Observe that the last row and last column of $\mathbf{G}_{CL}$ *corresponds to the cultural leader. We see that each of the four agents in the network is linked to him but he is only linked to himself* ($\mathbf{G}_{CL}$ *has a unique closed communication class).*

We now need to calculate the index given in equation (17) in Proposition 4 for each of the four followng cases: the government leader is connected to agent 1, to agent 2, to agent 3 and to agent 4. For the sake of the exposition, let us show how we calculate this index when *the government leader is connected to agent 1 since the calculations are similar for the three other cases.*

When the government leader $i = GL$ is connected to agent 1, the matrix of the network is now given by (where the last row and last column of this matrix corresponds to the government leader):

$$\mathbf{G}_{CL} + i1 = \left[\begin{array}{cccc|cc} 0 & \frac{1}{5} & \frac{1}{5} & \frac{1}{5} & \frac{1}{5} & \frac{1}{5} \\ \frac{1}{2} & 0 & 0 & 0 & \frac{1}{2} & 0 \\ 0 & \frac{1}{3} & 0 & \frac{1}{3} & \frac{1}{3} & 0 \\ 0 & 0 & \frac{1}{2} & 0 & \frac{1}{2} & 0 \\ \hline 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{array}\right]$$

In this matrix, we consider the top-left block that is given by the interactions just among the original agents in the new network. As stated above, we call this matrix $\mathbf{Q}_{CL} + i1$. It is equal to:

$$\mathbf{Q}_{CL} + i1 = \begin{bmatrix} 0 & \frac{1}{5} & \frac{1}{5} & \frac{1}{5} \\ \frac{1}{2} & 0 & 0 & 0 \\ 0 & \frac{1}{3} & 0 & \frac{1}{3} \\ 0 & 0 & \frac{1}{2} & 0 \end{bmatrix}$$

We then obtain:

$$\mathbf{M}(\mathbf{Q}_{CL} + i1) = \sum_{p=0}^{\infty} (\mathbf{Q}_{CL} + i1)^p = [\mathbf{I} - (\mathbf{Q}_{CL} + i1)]^{-1} = \begin{bmatrix} \frac{25}{21} & \frac{8}{21} & \frac{3}{7} & \frac{8}{21} \\ \frac{25}{42} & \frac{25}{21} & \frac{3}{14} & \frac{4}{21} \\ \frac{5}{21} & \frac{10}{21} & \frac{9}{7} & \frac{10}{21} \\ \frac{5}{42} & \frac{5}{21} & \frac{9}{14} & \frac{26}{21} \end{bmatrix}$$

We can now sum over the first column entries and divide by the degree of agent 1. We obtain: $\frac{1}{5}\left(\frac{25}{21} + \frac{25}{42} + \frac{5}{21} + \frac{5}{42}\right) = \frac{3}{7} = 0.4285$. *This is exactly the value of the index given in equation (17) in Proposition 4 when $k = 1$, i.e.* $0.4285 = \frac{1}{d_1} \sum_{l=1}^{4} m_{l1} \left(\mathbf{Q}_{CL} + i1\right)$, *where $i = GL$.*

We can perform exactly the same exercise when the government leader is connected to agent 2, 3 and 4. If we also add the value of the index obtained when the government leader is connected to agent 1 (i.e. 0.4285*), then we obtain the following vector:* $(0.4285, 0.7692, 0.6521, 0.7692)$. *We see that this index is the highest for agent 2 or 4 because they have the lowest outdegree but the highest Katz-Bonacich centrality in the augmented net-*

work where their link to the government leader is added. As a result, the government leader should be connected to agent 2 or 4 in the network because this maximizes total assimilation. Let us show that this is true in this example when we calculate the values of the sum of the long-run assimilation norms.

Let us now determine the long-run assimilation norms of the four agents in the network using the same method as in Section B.4 in the Online Appendix B. Recall that $s_{CL} = 0$ *and* $s_{GL} = 1$. *Since the two leaders are the only two closed communication classes, the long-run norms of all agents just depend on the norm values of the two leaders, independently of the original norms of the other agents. Following the results in* (10) *in Proposition 1, when we link the government leader to agent 1, we obtain:*

$$\mathbf{s}^{\infty}_{(\mathbf{G}_{CL}+i1)} = \lim_{t\to\infty} (\mathbf{G}_{CL} + i1)^t \, \mathbf{s}^{(0)}$$

which is equivalent to:

$$\mathbf{s}^{\infty}_{(\mathbf{G}_{CL}+i1)} = \left[\begin{array}{cccc|cc} 0 & 0 & 0 & 0 & \frac{16}{21} & \frac{5}{21} \\ 0 & 0 & 0 & 0 & \frac{37}{42} & \frac{5}{42} \\ 0 & 0 & 0 & 0 & \frac{20}{21} & \frac{1}{21} \\ 0 & 0 & 0 & 0 & \frac{41}{42} & \frac{1}{42} \\ \hline 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{array}\right] \left[\begin{array}{c} \mathbf{s}_1^{(0)} \\ \mathbf{s}_2^{(0)} \\ \mathbf{s}_3^{(0)} \\ \mathbf{s}_4^{(0)} \\ \hline 0 \\ 1 \end{array}\right]$$

As stated above, the initial norms of the four agents in the network are irrelevant because everything will depend on the norms of the stubborn agents. Solving this equation, we obtain:

$$\mathbf{s}^{\infty}\left(\mathbf{G}_{CL} + i1\right) = \begin{pmatrix} 0.24 \\ 0.12 \\ 0.05 \\ 0.02 \end{pmatrix}$$

where the aggregate assimilation level is: $\sum_{i=1}^{4} s_i^{\infty}(\mathbf{G}_{CL} + i1) = 0.43$.

We can do the same exercise by linking the government leader to agent 2, 3, and 4. It is *easily verifed that the vector of long-run norms in each case is given by:*

$$\mathbf{s}^{\infty}\left(\mathbf{G}_{CL}+i2\right)=\begin{pmatrix}0.15\\0.38\\0.15\\0.08\end{pmatrix};\ \mathbf{s}^{\infty}\left(\mathbf{G}_{CL}+i3\right)=\begin{pmatrix}0.13\\0.06\\0.30\\0.15\end{pmatrix};\ \mathbf{s}^{\infty}\left(\mathbf{G}_{CL}+i4\right)=\begin{pmatrix}0.15\\0.08\\0.15\\0.38\end{pmatrix}.$$

We easily obtain that: $\sum_{i=1}^{4} s_i^{\infty}(\mathbf{G}_{CL}+i2)=\sum_{i=1}^{4} s_i^{\infty}(\mathbf{G}_{CL}+i4)=0.76$ *and* $\sum_{i=1}^{4} s_i^{\infty}(\mathbf{G}_{CL}+i3)=0.64$*. As a result, the planner should link the government leader to agent 2 or agent 4 since it leads to the highest aggregate assimilation level in the network. Observe that, since we normalized the norms of the two leaders to 0 and 1, respectively, the value of the index and the value of the sum of aggregate assimilation norms are the same. For example, when the government leader connects to agent 2 or 4, the value of the index or the sum of aggregate assimilation norms are both equal to* 0.76*.*

3.4.3 Optimal network design

Let us go back to the benchmark model (without a cultural or a government leader) and let us determine the network that maximizes aggregate assimilation. Assume that the initial norms are given. Then, the policymaker would like to take the agent with the highest initial social norm and make him or her the only hub of the network. In other words, the optimal network for aggregate assimilation is a (directed) *star network* in which everyone is linked to the agent with the highest initial social norm and this agent is connected to nobody. In that case, there will be a unique close communication class, the star, and all the other agents (including the star) will converge to the social norm of the star.

3.5 Speed of Convergence

In the previous sections, we focused on the convergence levels of assimilation norms. We now analyze the *speed of convergence* of norms. For policy purpose, assimilation is important only if it is reached reasonably quickly. If it takes many years for individuals to be assimilated, then the model's limit predictions are unlikely to be useful.

In Proposition 10 in the Online Appendix D, we show that, depending on the taste for conformity ω and on the parameter γ, the convergence of social norms can be faster in our model than under the standard updating rule (Golub and Jackson (2010)) using the matrix $\mathbf{G}$. In the Online Appendix D, we also provide some simulations illustrating this result.

3.6 Definition of peers

So far, we have defined peers as *outdegrees*, i.e. there is a link ($g_{ij} = 1$) between two individuals i and j if individual i nominates individual j. It did not matter if the reverse was true. Furthermore, if individual j nominates individual i but not the reverse, then $g_{ji} = 1$. In empirical research, peers can be defined in different ways: they can be defined as outdegrees, or indegrees or both or either. For example, in empirical studies, researchers have extensively been using the National Longitudinal Survey of Adolescent Health (AddHealth), which has information on friendship networks. Using the AddHealth data, Calvó-Armengol et al. (2009) define a link between two persons whenever one of them has nominated the other. In our model, this would imply that the network will be strongly connected. Using the same dataset, Haynie (2001) and Lin (2010) define peers as outdegrees. This is the way we have defined peers so far. Other studies, such as Laursen (1993) and Erwin (1998) argue that friendships are reciprocal by definition so that both persons need to nominate each other for a link in the network to exist.

Obviously, the way links or peers are defined has an important impact on the definition of the closed communication classes in a network and, thus, on the convergence of social norms and actions of all individuals in a network. Thus, in the Online Appendix E, we consider two other definitions of peers widely used in the empirical literature on peer and network effects,[17] that is peers as *indegrees* and peers as *mutual friends*, and show how these definitions have a strong impact on the long-run choices of social norms and actions.

4 The Economic Incentives for Assimilation

4.1 Model and steady-state equilibrium

In this section, we extend our benchmark model to include the idiosyncratic incentives of agents for assimilation. Each agent, in fact, depending on her assimilation choices, may have different gains from assimilation in terms of job opportunities, salary, and other economically relevant outcome. We do not model here how differential incentives are produced, but we just assume that each agent has a particular incentive to perform a particular assimilation choice, apart from own norms and conformism motives. This section explores how results change when this is taken into account and how the policy maker can exploit this new aspect of assimilation.

[17] See, Boucher and Fortin (2016), Chandrasekhar (2016), Advani and Malde (2018) for overviews of this literature.

In what follows, we propose a more general utility function which, together with the losses associated to conformism and consistency, also allows for some idiosyncratic gains of the action. There are indeed economic gains of assimilation and these gains are heterogeneous, depend on individual characteristics and do not vary over time. For that, we introduce some ex ante heterogeneity (i.e. observable characteristics such as gender, race, parental education, etc.) in the utility function, which is now defined as:

$$u_i^t = \underbrace{2\alpha_i x_i^t - \left(x_i^t\right)^2}_{\text{Idiosyncratic term}} \underbrace{- \omega(x_i^t - \sum_j g_{ij} x_j^t)^2}_{\text{Conformism}} - \underbrace{(x_i^t - s_i^t)^2}_{\text{Consistency}} \tag{18}$$

where $\alpha_i > 0$ captures the ex ante heterogeneity of agent i. The higher is $\alpha_i > 0$, the greater is the individual marginal utility of exerting action x_i.[18] The coefficient α_i is what we call economic incentive to assimilation and can be altered by targeted policies aiming at making assimilation more or less profitable to agents. First-order condition now yields:

$$x_i^t = \left(\frac{1}{2+\omega}\right)\alpha_i + \left(\frac{1}{2+\omega}\right)s_i^t + \left(\frac{\omega}{2+\omega}\right)\sum_j g_{ij} x_j^t \tag{19}$$

The equilibrium actions are now a convex linear combination of the individual i's marginal incentives, α_i, her own norm, s_i^t and the average actions of her neighbors, $\sum_j g_{ij} x_j^t$. Denote $\theta' \equiv \omega/(2+\omega)$. We then obtain the equilibrium actions in matrix form:

$$\mathbf{x}^t = \left(\frac{1}{2+\omega}\right)\left[\mathbf{I} - \theta'\,\mathbf{G}\right]^{-1}\left(\mathbf{s}^t + \boldsymbol{\alpha}\right) \tag{20}$$

where $\boldsymbol{\alpha}$ is the vector of α_i. While this formulation is relatively simple, we cannot go further in the characterization of the dynamics and the equilibrium behavior in terms of the network $\mathbf{G}$ because we cannot express the dynamics as a linear system $\mathbf{x}^t = \mathbf{A}\mathbf{x}^{t-1}$ as before.

Therefore, we propose an approach that allows us to solve this problem in a simple way. For this purpose, we define by $\hat{\mathbf{G}}$ an *augmented* $\mathbf{G}$ matrix, where we consider a fictitious directed network of $2n$ agents where, on top of network $\mathbf{G}$, for each agent i, we create a *fictitious* agent f_i. Each agent i is linked to f_i, no other agent in the network is linked to f_i, and f_i has no other link than i but herself (self-loop). Call F the set of fictitious agents. Formally $\hat{g}_{i,f_i} > 0$, $\hat{g}_{f_i,j} = 0$ and $\hat{g}_{f_i,f_i} = 1$. In particular, the first n rows of $\hat{\mathbf{G}}$ represent the links of fictitious agents while the last n rows represent the links of real agents. Remembering

[18] Note that we have $2\alpha_i$ instead of α_i just to ease computations, and this is without loss of generality.

that $\theta = \omega/(1+\omega)$, we can define the *augmented* $\mathbf{G}$ matrix in its canonical form as follows:

$$\hat{\mathbf{G}} \equiv \left[\begin{array}{c|c} \mathbf{I} & \mathbf{O} \\ \hline \mathbf{D} & \theta\mathbf{G} \end{array}\right] \tag{21}$$

where each block is of dimension $n \times n$, where $\mathbf{D}$ is a diagonal matrix with entries $1/(1+\omega)$ and $\mathbf{O}$ is a matrix of zeros. Observe that $\hat{\mathbf{G}}$ is a stochastic matrix. We study now the optimal action profile $\hat{\mathbf{x}}^t$, in which we assume that each fictitious agent just plays own α_i. Then, the first-order condition (19) can be written as follows:

$$\hat{\mathbf{x}}^t = \left(\frac{1}{2+\omega}\right)\hat{\mathbf{s}}^t + \left(\frac{1+\omega}{2+\omega}\right)\hat{\mathbf{G}}\hat{\mathbf{x}}^t \tag{22}$$

where

$$\hat{\mathbf{x}}^t = \left[\begin{array}{c} \boldsymbol{\alpha} \\ \hline \mathbf{x}^t \end{array}\right], \qquad \hat{\mathbf{s}}^t = \left[\begin{array}{c} \boldsymbol{\alpha} \\ \hline \mathbf{s}^t \end{array}\right] \tag{23}$$

This implies that (20) can now be written as:

$$\hat{\mathbf{x}}^t = \left(\frac{1}{2+\omega}\right)\left[\mathbf{I} - \left(\frac{1+\omega}{2+\omega}\right)\hat{\mathbf{G}}\right]^{-1}\hat{\mathbf{s}}^t = \left(\frac{1}{2+\omega}\right)\left[\mathbf{I} - \frac{\theta'}{\theta}\hat{\mathbf{G}}\right]^{-1}\hat{\mathbf{s}}^t \tag{24}$$

Instead of working with $\mathbf{G}$, we will now work with $\hat{\mathbf{G}}$, being aware that the first n rows of all matrices and vectors are about fictitious agents and, thus, without economic meaning. In these cases, let us show that all the previous results hold. Define $\hat{\mathbf{T}}$ as the augmented $\mathbf{T}$ matrix (i.e. by adding the n fictitious players). We also have:

$$\hat{\mathbf{s}}^{t+1} = \gamma \underbrace{\hat{\mathbf{x}}^t}_{\text{Consistency}} + (1-\gamma)\underbrace{\hat{\mathbf{s}}^t}_{\text{Anchoring}} \tag{25}$$

Plugging the equilibrium action in (24), and calling $\mathbf{M}(\frac{\theta'}{\theta}, \hat{\mathbf{G}}) := \left[\mathbf{I} - \frac{\theta'}{\theta}\hat{\mathbf{G}}\right]^{-1}$, we obtain:

$$\hat{\mathbf{s}}^{t+1} = \left[\frac{\gamma}{2+\omega}\mathbf{M}(\frac{\theta'}{\theta}, \hat{\mathbf{G}}) + (1-\gamma)\mathbf{I}\right]\hat{\mathbf{s}}^t \tag{26}$$

so that $\hat{\mathbf{T}} := \frac{\gamma}{2+\omega}\mathbf{M}(\frac{\theta'}{\theta}, \hat{\mathbf{G}}) + (1-\gamma)\mathbf{I}$. It is straightforward to show that:

Corollary 1 *For any* $\mathbf{G}$*, we have:*

$$\hat{\mathbf{s}}^{\infty} = \lim_{t\to\infty} \hat{\mathbf{T}}^t\, \hat{\mathbf{s}}^{(0)} = \lim_{t\to\infty} \hat{\mathbf{G}}^t\, \hat{\mathbf{s}}^{(0)} \tag{27}$$

Observe that, in network $\hat{\mathbf{G}}$, each node f_i forms a closed communication class and there are no other closed communication classes. Formally, $C_{f_i} = \{f_i\}$ for all i. Therefore, the dynamics of the norm (and action) is now represented by a matrix with n closed communication classes and n agents belonging to other communication classes. We can thus use the results of Proposition 9 of the Online Appendix B.4 to obtain:

Proposition **5** *If the utility of each agent i is given by (18), then the steady-state norms of all n agents in the original network are given by:*

$$\mathbf{s}^{\infty} = \left(\frac{1}{1+\omega}\right)[\mathbf{I} - \theta\mathbf{G}]^{-1}\boldsymbol{\alpha} \tag{28}$$

while the steady-state actions are equal to:

$$\mathbf{x}^{\infty} = \left(\frac{1}{2+\omega}\right)[\mathbf{I} - \theta'\,\mathbf{G}]^{-1}\left[\mathbf{I} + \frac{1}{(1+\omega)}(\mathbf{I} - \theta\mathbf{G})^{-1}\right]\boldsymbol{\alpha} \tag{29}$$

Indeed, independently of the initial norms, in the long run, assimilation norms and choices are totally determined by the ex ante heterogeneities of agents in terms of α_i and their position of the network. Since $\boldsymbol{\alpha}$ can also represent a vector of heterogenous economic incentives for the agents, this result means that economic incentives can drive social norms very far from their initial range. Suppose, for example, that the initial norms are such that $\mathbf{s}^{(\mathbf{0})} \in [0,1]^n$, but that $\boldsymbol{\alpha} \in [2,3]^n$, then $\mathbf{s}^{(\infty)} \in [2,3]^n$. In this respect, initial norms become irrelevant and updated norms and actions only follow the economic incentives. Finally, it is worth noticing that the way in which agents interact in imitating norms determines in a precise manner the final distribution of norms, and that the weights are proportional to the Katz-Bonacich centrality of agents in the $\mathbf{G}$ network.

Example **3** *To illustrate these results, consider the network with 4 agents given in Figure 1 and let us determine the augmented network. The augmented* $\mathbf{G}$ *with the four fictitious* *agents is given by:*

$$\hat{\mathbf{G}} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ \frac{1}{1+\omega} & 0 & 0 & 0 & 0 & \frac{\omega}{3(1+\omega)} & \frac{\omega}{3(1+\omega)} & \frac{\omega}{3(1+\omega)} \\ 0 & \frac{1}{1+\omega} & 0 & 0 & \frac{\omega}{1+\omega} & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{1+\omega} & 0 & 0 & \frac{\omega}{2(1+\omega)} & 0 & \frac{\omega}{2(1+\omega} \\ 0 & 0 & 0 & \frac{1}{1+\omega} & 0 & 0 & \frac{\omega}{1+\omega} & 0 \end{bmatrix} \tag{30}$$

and the corresponding network can be displayed as follows:

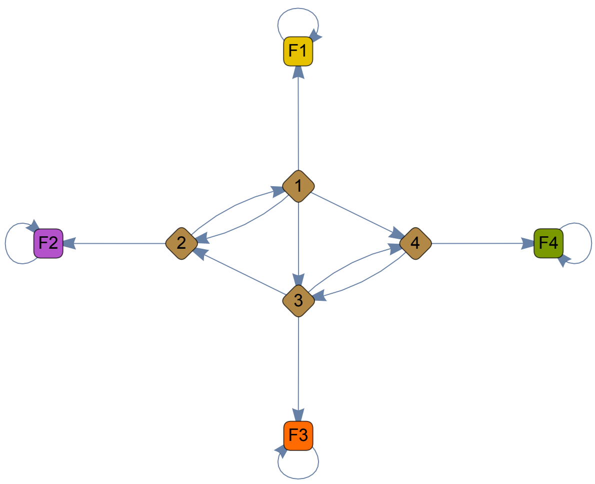

Figure 5: The *augmented* $\mathbf{G}$ network

Take

$$\boldsymbol{\alpha} = \begin{bmatrix} 0.1 \\ 0.2 \\ 0.3 \\ 0.4 \end{bmatrix}, \qquad \mathbf{s}^{(0)} = \begin{bmatrix} 0.6 \\ 0.7 \\ 0.8 \\ 0.9 \end{bmatrix}$$

Consider now the case of $\omega = 0.5$*. Then*

$$\mathbf{M}(\frac{\theta'}{\theta},\hat{\mathbf{G}}) = \begin{bmatrix} 2.5 & 0. & 0. & 0. & 0. & 0. & 0. & 0. \\ 0 & 2.5 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2.5 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2.5 & 0 & 0 & 0 & 0 \\ 1.015 & 0.076 & 0.083 & 0.076 & 1.015 & 0.076 & 0.0839 & 0.076 \\ 0.203 & 1.015 & 0.016 & 0.015 & 0.203 & 1.015 & 0.016 & 0.015 \\ 0.0202 & 0.103 & 1.022 & 0.103 & 0.020 & 0.103 & 1.022 & 0.103 \\ 0.004 & 0.0202 & 0.204 & 1.020 & 0.004 & 0.020 & 0.204 & 1.020 \end{bmatrix}$$

Setting $\gamma = 0.5$ *we have that simply* $\hat{\mathbf{T}} = 0.2\mathbf{M}(\frac{\theta'}{\theta},\hat{\mathbf{G}}) + 0.5\mathbf{I}$. *We also have:*

$$\mathbf{T}^{\infty} = \hat{\mathbf{G}}^{\infty} = \begin{bmatrix} 1. & 0. & 0. & 0. & 0. & 0. & 0. & 0. \\ 0. & 1. & 0. & 0. & 0. & 0. & 0. & 0. \\ 0. & 0. & 1. & 0. & 0. & 0. & 0. & 0. \\ 0. & 0. & 0. & 1. & 0. & 0. & 0. & 0. \\ 0.699 & 0.0964 & 0.109 & 0.096 & 0. & 0. & 0. & 0. \\ 0.233 & 0.699 & 0.036 & 0.032 & 0. & 0. & 0. & 0. \\ 0.041 & 0.123 & 0.712 & 0.123 & 0. & 0. & 0. & 0. \\ 0.014 & 0.041 & 0.237 & 0.708 & 0. & 0. & 0. & 0. \end{bmatrix}$$

The long-run values of the assimilation norms and efforts for the "real" agents are given by:

$$\mathbf{s}^{\infty} = \begin{bmatrix} 0.699 & 0.0964 & 0.109 & 0.096 \\ 0.233 & 0.699 & 0.036 & 0.032 \\ 0.041 & 0.123 & 0.712 & 0.123 \\ 0.014 & 0.041 & 0.237 & 0.708 \end{bmatrix} \begin{bmatrix} 0.6 \\ 0.7 \\ 0.8 \\ 0.9 \end{bmatrix} = \begin{bmatrix} 0.160 \\ 0.187 \\ 0.292 \\ 0.364 \end{bmatrix}, \qquad \mathbf{x}^{\infty} = \begin{bmatrix} 0.369 \\ 0.368 \\ 0.498 \\ 0.579 \end{bmatrix}$$

We can see that, even if all four agents inherited social norms favorable to assimilation, they end up having norms and efforts that are not in favor of assimilation. This is because their economic incentives of assimilation (i.e. their ex ante observable characteristics) were very different from their initial norms and less conducive to assimilation. Note that individual 1, *who is the most central person in the network but has a very low incentive to assimilate, end up having the lowest assimilation norm and effort.*

4.2 Policies

We have shown that the long-run norms and actions are now totally driven by the economic incentives for assimilation of agents. We now analyze how a planner can provide incentives to each individual in order to reach a certain profile of long-term norms and thus assimilation choices.[19] Define $\mathbf{s}^*$, $\boldsymbol{\alpha^*}$, and $\boldsymbol{\sigma^*}$ to be, respectively, the target vector of long-term norms, the incentives vector in terms of the αs that enables these norms to be reached, and $\boldsymbol{\sigma^*} := \boldsymbol{\alpha^*} - \boldsymbol{\alpha}$. Then $\boldsymbol{\sigma^*}$ is the vector of additional positive (subsidies) or negative (taxes) incentives that the planner should provide to each agent in order to reach $\mathbf{s}^*$. In terms of timing, we just add one stage at time $t = 0$ where the planner set the incentives $\boldsymbol{\sigma^*}$ at each period of time.

PROPOSITION **6** *In order to reach the long-term norms* $\mathbf{s}^*$*, the planner must give the following incentives to each individual:*

1. *If the objective is to reach the same long-term norm for all individuals, i.e.* $s_i^* = s_j^*$ *for all* i, j*, then* $\boldsymbol{\sigma^*} = \boldsymbol{s^*} - \boldsymbol{\alpha}$*;*

2. *If the objective is to have different long-term norms for different individuals, i.e.* $s_i^* \neq s_j^*$ *for all* i, j*, then* $\boldsymbol{\sigma^*} = \mathbf{s}^* - \boldsymbol{\alpha} + \omega(\mathbf{I} - \mathbf{G})\mathbf{s}^*$*.*

If the policymaker is interested in having an homogeneous society where all individuals reach the same (assimilation) norm in the long run, then the marginal subsidy must be the difference between the desired norm and the actual economic incentive, irrespectively of the network. Intuitively, $\boldsymbol{\sigma^*}$ make all agents ex ante equal in terms of incentives. Since long-run norms are a convex combination of the $\boldsymbol{\alpha^*}$, it is straightforward to see that the norms will converge to the same value.

If the planner would like to have an heterogenous distribution of long-term norms, then it has to take into account the direct effect of subsidies on each individual but also on their neighbors. In Proposition 6, we show the per-person subsidy depends on the taste for conformity, the network structure and the ex ante heterogeneity in terms of αs.

Indeed, by definition of $\boldsymbol{\sigma^*}$, we first notice that:

$$\boldsymbol{\alpha^*} = (1 + \omega)(\mathbf{I} - \theta\mathbf{G})\mathbf{s}^* \tag{31}$$

In this equation, $(\mathbf{I} - \theta\mathbf{G})\mathbf{s}^*$ tells us how much each s_i^* differs form the average (the social norm). Moreover, in the dynamics, each agent is affected by others depending on the parameter ω. Then the incentives α_i should be equal to the desired norm, corrected for how

[19] Contrary to Section 3.4.2, we do not consider either a cultural or a government leader because stubborn agents here have no impact on the assimilation process of agents in the network.

much the desired norm is larger or smaller than the average desired norm, corrected for the level of conformism.

Example 4 *To get an intuition about how this policy is shaped by the network structure and the desired distribution of long-term norms, consider the networks displayed in Figure 3 and the following vector of ex ante heterogeneities:*

$$(\boldsymbol{\alpha})^T = \begin{pmatrix} 0.1 & 0.2 & 0.3 & 0.4 & 0.5 & 0.6 & 0.7 & 0.8 & 0.9 \end{pmatrix} \tag{32}$$

Assume first that the planner wants an homogenous society with a maximal level of integration, i.e. $s_i^* = 1$ *for all* i*. Then,* $\boldsymbol{\sigma}^* = \boldsymbol{s}^* - \boldsymbol{\alpha}$*. Thus,*

$$(\boldsymbol{\sigma})^T = \begin{pmatrix} 0.9 & 0.8 & 0.7 & 0.6 & 0.5 & 0.4 & 0.3 & 0.2 & 0.1 \end{pmatrix} \tag{33}$$

In steady state, everyone will have a social norm of $s_i^* = 1$ *and will play* $x_i^* = 1$*. There is thus total assimilation. This is because, for every individual,* $\alpha_i^* = \alpha_i + \sigma_i^* = 1$*, and actions are a convex linear combinations of all* α^**s. In that case, the overall cost of the subsidy is given by:* $\sum_i \sigma_i^* x_i^* = \sum_i \sigma_i$*, independently of the network structure. Thus, if the desired norm is the same for everyone, the determination of the subsidy and its cost is independent of the network structure.*

Consider, now, the case where the planner has different targets in terms of social norms. For example, consider the case in which the policymaker does not want any agent to have a long-term norm of assimilation below 0.5*. Consider the first two networks in Figure 3 and the ex ante heterogeneity is given by* (32)*. If there no intervention, the long-run norms in each network are respectively given by:*[20]

$$\mathbf{s}_{\mathbf{G}_1}^{\infty} = \begin{pmatrix} 0.201 & 0.223 & 0.270 & 0.450 & 0.539 & 0.565 & 0.669 & 0.785 & 0.842 \end{pmatrix} \tag{34}$$

$$\mathbf{s}_{\mathbf{G}_2}^{\infty} = \begin{pmatrix} 0.154 & 0.220 & 0.262 & 0.450 & 0.539 & 0.565 & 0.664 & 0.784 & 0.841 \end{pmatrix} \tag{35}$$

Consider, now, the policy above where the objective is that everybody needs to have a level of long-term assimilation of at least 0.5*. This means that, in the steady state with no intervention, the individuals who have an assimilation value below* 0.5 *need to increase it to* 0.5 *while those with values above or equal to* 0.5 *keep this same level. This implies that, for the*

[20] Even if network 1 is strongly connected, the individuals will not have the same long-term assimilation norm because they have different αs. The same reasoning applies for network 2, which has one closed communication class.

first two networks in Figure 3, the objective is:

$$\mathbf{s}^*_{\mathbf{G}_1} = \begin{pmatrix} 0.5 & 0.5 & 0.5 & 0.5 & 0.539 & 0.565 & 0.669 & 0.785 & 0.842 \end{pmatrix} \tag{36}$$

$$\mathbf{s}^*_{\mathbf{G}_2} = \begin{pmatrix} 0.5 & 0.5 & 0.5 & 0.5 & 0.539 & 0.565 & 0.664 & 0.784 & 0.841 \end{pmatrix} \tag{37}$$

Then, using (31)*, we can compute the subsidy given to each agent in order to reach these steady-state norms:*

$$\boldsymbol{\alpha}^*_{\mathbf{G}_1} = \begin{pmatrix} 0.490 & 0.5 & 0.5 & 0.473 & 0.491 & 0.587 & 0.650 & 0.8 & 0.9 \end{pmatrix} \tag{38}$$

$$\boldsymbol{\alpha}^*_{\mathbf{G}_2} = \begin{pmatrix} 0.5 & 0.5 & 0.5 & 0.473 & 0.491 & 0.587 & 0.642 & 0.8 & 0.9 \end{pmatrix} \tag{39}$$

Notice that, in the second network, the first three agents need an incentive equal to the targeted norm. This is because they form a closed communication class, and thus if the policymaker wants them to reach at least 0.5*, she has to induce this behavior through incentives. Interestingly, agents 4 and 5 need incentives lower than* 0.5 *since imitation and conformism lead them to have higher long-run norms. As a result, the final marginal subsidy or tax for each individual should be as follows:*

$$\boldsymbol{\sigma}^*_{\mathbf{G}_1} = \begin{pmatrix} 0.390 & 0.3 & 0.2 & 0.073 & -0.008 & -0.012 & -0.049 & -1.110 \times 10^{-16} & -1.1102 \times 10^{-16} \end{pmatrix}$$

$$\boldsymbol{\sigma}^*_{\mathbf{G}_2} = \begin{pmatrix} 0.4 & 0.3 & 0.2 & 0.073 & -0.008 & -0.012 & -0.057 & 1.110 \times 10^{-16} & -1.110 \times 10^{-16} \end{pmatrix}$$

The policymaker gives a marginal subsidy to the first four agents and imposes a tax to all the other agents. Note, however, that we did not impose any budget constraint so that the cost of the policy can be positive or negative depending on the parameters. Indeed, the cost of the policy is given by $(\boldsymbol{\sigma}^*)^T\mathbf{x}$ *and, in this example, is equal to* 0.85 *(network 1) and* 0.91 *(network 2).*

Given that the objective is to have all agents converging to a steady-state social norm at least equal to 0.5*, the cost would only be reduced by lowering the assimilation norms and the choices of agents above* 0.5 *in equilibrium. In the example, the cost would be set to 0 once we choose an homogeneous steady-state norm vector of* 0.5*. In this case, everyone would choose an action equal to* 0.5*. Then, the agent with the initial norm of* 0.9 *will pay for the subsidy of the agent with an initial norm of* 0.1*. The agent with initial norm of* 0.8 *would pay for the agent with an initial norm of* 0.2*, and so forth. In this case, the overall cost of this policy is zero and thus this policy is self-financed.*

5 Other applications

So far, the model has been interpreted in terms of assimilation norms. It can have many other applications. For example, another natural application of our framework is to assume that $x_i^t = 1 - y_i^t$ is the effort in crime while y_i^t is the effort exerted in labor, with these two activities being perfectly substitutables. It is indeed well-documented that being a criminal and a worker at the same time is quite common (see e.g. Freeman (1996)). For example, drug dealers often hold low-skilled jobs. This implies that the higher an individual i exerts effort in crime, the lower she spends time working. In that interpretation, all our results go through and we can explain how social norms, network position and ex ante heterogeneity affect crime and labor. In particular, we can show that if some individuals live in segregated communities, isolated from the rest of the society, then they will mostly commit crimes if their individual abilities and their individual norms (such as work ethics) are not in favor of working in the labor market.

Another straighforward application is tax evasion. There is plenty of evidence showing that social norms and social interactions matter in the decision of tax evading (see e.g. Andreoni et al. (1998), Posner (2000) Fortin et al. (2007), Luttmer and Singhal (2014), Besley et al. (2014)). Indeed, when deciding to evade, social norms play a crucial role in guilt and shame in tax compliance behavior and, as argued by Gordon (1989) and Myles and Naylor (1996), each individual can derive a psychic payoff from adhering to the standard pattern of reporting behavior in her reference group (social conformity effect). As noted by Luttmer and Singhal (2014), cultural or social norms can affect the strength of the individual's intrinsic motivations to pay taxes or the sensitivity to peers. If we interpret x_i^t as the fraction of income that is evaded at time t, then we can use our model to explain how the tax compliance behavior is affected by social norms, intrinsic motivations and initial "honesty" norms. Our model shows that individuals will be more likely not to evade and to pay taxes if they have intrinsic motivation to pay taxes or feel guilt or shame for failure to comply and if they may be influenced by peer behavior and the possibility of social recognition or sanctions from peers. Different tax compliance behaviors may then arise in the long-run equilibrium depending on these different aspects, especially the different social norms that exist in each group or community or even country. For example, according to the IMF estimates, 30 percent of taxes were evaded in Greece in 2011 while the same number was 7 percent in the UK (IMF (2013)). According to this report, the main explanation put forward for this gap is the difference in social norms between these two countries.[21]

[21] Acemoglu and Jackson (2017) propose an interesting but different model based on social norms and the enforcement of laws that can explain why Greece and the UK experienced different tax evasion rates.

6 Conclusion

We consider a model where each individual (or ethnic minority) is embedded in a network of relationships and decides whether or not she wants to be assimilated to the majority norm. First, each individual wants her behavior to agree with her personal ideal action or norm, which means that there is a *consistency* between her own norm and her assimilation behavior. Second, she also wants her behavior to be as close as possible to the average assimilation behavior of her peers, which implies that she is a *conformist*. We show that there is always convergence to a steady-state and characterize it. We also derive some implications in terms of assimilation policies. More generally, our model highlights the tension that may exist between the social norms of each family, that of the community and own marginal benefits in the long-run assimilation decisions of ethnic minorities.

We view our model as a step forward an analysis of how pressure from peers, communities and families affect the long-run decisions of individuals, especially for decisions on assimilation, crime, tax evasion, etc. New steps in this research should incorporate an empirical analysis of these interactions in order to disentangle the different effects at work and to address the relevant policy implications of such an analysis.

References

Acemoglu, D. and M.O. Jackson (2017), "Social norms and the enforcement of laws," *Journal of the European Economic Association* 15, 245–295.

Advani, A. and B. Malde (2018), "Methods to identify linear network models: A review," *Swiss Journal of Economics and Statistics*, forthcoming.

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.

Akerlof, G.A. (1997), "Social distance and social decisions," *Econometrica* 65, 1005–1027.

Akerlof, G.A. and W. Dickens (1982), "The economic consequences of cognitive dissonance" *American Economic Review* 72, 307–319.

Alba, R. and V. Nee (1997), "Rethinking assimilation theory for a new era of immigration," *International Migration Review* 31, 826-?874.

Algan, Y., Bisin, A., Manning, A. and T. Verdier (2012), *Cultural Integration of Immigrants in Europe*, Oxford: Oxford University Press.

Andreoni, J., Erard, B. and J. Feinstein (1998), "Tax compliance," *Journal of Economic Literature* 36, 818–860.

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

Bem, D. (1972), "Self-perception theory," In: L. Berkowitz (Ed.), *Advances in Experimental Social Psychology*, New York: Academic Press, pp. 1–62.

Benabou, R. and J. Tirole (2004), "Will power and personal rules," *Journal of Political Economy* 112, 848–887.

Benabou, R. and J. Tirole (2006), "Incentives and prosocial behavior," *American Economic Review* 96, 1652–1678.

Benhabib, J., Bisin, A. and M.O. Jackson (2011), *Handbook of Social Economics* Vol. 1 and 2, Amsterdam: Elsevier Science.

Beenstock, M., Chiswick, B.R. and A. Paltiel (2010), "Testing the immigrant assimilation hypothesis with longitudinal data," *Review of Economics of the Household* 8, 7–27.

Bertrand, M., Luttmer, E.F.P., and S. Mullainathan (2000), "Network effects and welfare cultures," *Quarterly Journal of Economics* 115, 1019–1055.

Besley, T., Jensen, A. and T. Persson (2014), "Norms, enforcement, and tax evasion," Unpublished manuscript, London School of Economics.

Bisin, A., Patacchini, E., Verdier, T. and Y. Zenou (2008), "Are Muslim immigrants different in terms of cultural integration?" *Journal of the European Economic Association* 6, 445–456.

Bisin, A. and T. Verdier (2000), "Beyond the melting pot: Cultural transmission, marriage, and the evolution of ethnic and religious traits," *Quarterly Journal of Economics* 115, 955–988.

Bisin, A. and T. Verdier (2001), "The economics of cultural transmission and the dynamics of preferences,," *Journal of Economic Theory* 97, 298–319.

Bonacich P. (1987), “Power and centrality: A family of measures,” *American Journal of Sociology* 92, 1170–1182.

Borjas, G.J. (1985), “Assimilation, changes in cohort quality, and the earnings of immigrants,” *Journal of Labor Economics* 3, 463–489.

Borjas, G.J. (1987), “Self-selection and the earnings of immigrants,” *American Economic Review* 77, 531–553.

Boucher, V. (2016), “Conformism and self-selection in social networks,” *Journal of Public Economics* 136, 30–44.

Boucher, V. and B. Fortin (2016), “Some challenges in the empirics of the effects of networks,” in *The Oxford Handbook on the Economics of Networks*, edited by Bramoullé, Yann, Andrea Galeotti and Brian Rogers, Oxford: Oxford University Press.

Bramoullé, Y., Rogers, B.W. and A. Galeotti (2016), *The Oxford Handbook of the Economics of Networks*, Oxford: Oxford University Press.

Brubaker, R. (2001), “The return of assimilation? Changing perspectives on immigration and its sequels in France, Germany, and the United States,” *Ethnic and Racial Studies* 24, 531–548.

Brueckner, J.K. and O. Smirnov (2007), “Workings of the melting pot: Social networks and the evolution of population attributes,” *Journal of Regional Science* 47, 209–228.

Brueckner, J.K. and O. Smirnov (2008), “Social networks and the convergence of population attributes: A generalization,” *Journal of Regional Science* 48, 359–365.

Buechel, B., Hellmann, T. and M.M. Pichler (2014), “The dynamics of continuous cultural traits in social networks,” *Journal of Economic Theory* 154, 274–309.

Buechel, B., Hellmann, T. and S. Klöï¿½ner (2015), “Opinion dynamics and wisdom under conformity,” *Journal of Economic Dynamics and Control* 52, 240–257.

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

Carvalho, J.-P. and M. Koyama (2016), “Jewish emancipation and schism: Economic development and religious change,” *Journal of Comparative Economics* 44, 562–584.

Chandrasekhar, A.G. (2016), “Econometrics of network formation,” in *The Oxford Handbook on the Economics of Networks*, edited by Bramoullé, Yann, Andrea Galeotti and Brian Rogers, Oxford: Oxford University Press.

Chandrasekhar, A.G., Larreguy, H. and J.P. Xandri (2015), “Testing models of social learning on networks: Evidence from a lab experiment in the field,” NBER Working Paper No. 21468.

Chiswick, B.R. and P.W. Miller (2002), “Immigrant earnings: Language skills, linguistic concentrations and the business cycle,” *Journal of Population Economics* 15, 31–57.

Chiswick, B.R. and P.W. Miller (2011), “The ‘negative’ assimilation of immigrants: A special case,” *Industrial and Labor Relations Review* 64, 502-525.

Corazzini, L., Pavesi, F., Petrovich, B. and L. Stanca (2012), “Influential listeners: An experiment on persuasion bias in social networks,” *European Economic Review* 56, 1276–1288.

Damm, A.P. (2009), “Ethnic Enclaves and Immigrant Labor Market Outcomes: Quasi-Experimental Evidence,” *Journal of Labor Economics* 27, 281–314.

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

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

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

Edin, P.-A., Fredriksson, P. and O. Åslund (2003), “Ethnic enclaves and the economic success of immigrants. Evidence from a natural experiment,” *Quarterly Journal of Economics* 118, 329–57.

Erwin, P. (1998), *Friendship in Childhood and Adolescence*, New York: Routledge.

Fordham, S. and J. Ogbu (1986), “Blacks students’ school successes: Coping with the burden of ‘acting white’,” *The Urban Review* 18, 176–206.

Fortin, B., Lacroix, G. and M.-C. Villeval (2007), “Tax evasion and social interactions,” *Journal of Public Economics* 91, 2089–2112.

Freeman, R.B. (1996), "Why do so many young American men commit crime and what can we do about it?" *Journal of Economic Perspectives* 10, 25–42.

Friedkin, N.E. and E.C. Johnsen (1999), "Social influence networks and opinion change," *Advances in Group Processes* 16, 1–29.

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

Furtado, D. and N. Theodoropoulos (2010), "Why does intermarriage increase immigrant employment? The role of networks," *B.E. Journal of Economic Analysis and Policy* 10(3), Article 101.

Gang, I.N. and K.F. Zimmermann (2000), "Is child like parent? Educational attainment and ethnic origin," *Journal of Human Resources* 35, 550–569.

Golub, B. and M.O. Jackson (2010), "Naive learning in social networks and the wisdom of crowds," *American Economic Journal: Microeconomics* 2, 112–149.

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

Golub, B. and E. Sadler (2016), "Learning in social networks," In: Bramoullé, Y., A. Galeotti and B.W. Rogers (Eds.), *The Oxford Handbook on the Economics of Networks*, Oxford: Oxford University Press.

Gordon, J.P.F. (1989), "Individual morality and reputation costs as deterrents to tax evasion," *European Economic Review* 33, 797-?805.

Hauk, E. and H. Mueller (2015), "Cultural leaders and the clash of civilizations," *Journal of Conflict Resolution* 59, 367–400.

Haynie, D.L. (2001), "Delinquent peers revisited: Does network structure matter?" *American Journal of Sociology* 106, 1013–1057.

Ioannides, Y.M. (2012), *From Neighborhoods to Nations: The Economics of Social Interactions*, Princeton: Princeton University Press.

IMF (2013), "Greece: IMF Country Report No. 13/155," Washington DC.

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

Jackson, M.O. (2014), "Networks in the understanding of economic behaviors," *Journal of Economic Perspectives*, 28(4), 3–22.

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

Jackson, M.O., Rogers, B.W. and Y. Zenou (2017), "The economic consequences of social network structure," *Journal of Economic Literature* 55, 49- -95.

Katz, L. (1953), "A new status index derived from sociometric analysis," *Psychometrika* 18, 39–43.

Kahanec, M. and K.F. Zimmermann (2011), *Ethnic Diversity in European Labor Market. Challenges and Solutions*, Cheltenham: Edward Elgar.

Kuran, T. and W.H.Sandholm (2008), "Cultural integration and its discontents," *Review of Economic Studies* 75, 201–228

Laursen, B. (1993), *Close Friendships in Adolescence*, San Francisco: Jossey-Bass.

Lin, X. (2010), "Identifying peer effects in student academic achievement by spatial autoregressive models with group unobservables," *Journal of Labor Economics* 28, 825–860.

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

Luttmer, E.F.P. and M. Singhal (2014), "Tax morale," *Journal of Economic Perspectives* 28, 149?-168.

Maani, S.A., Wang, X. and A. Rogers (2015), "Network effects, ethnic capital and immigrants' earnings assimilation: Evidence from a spatial, Hausman-Taylor estimation," IZA Discussion Paper No. 9308.

Meyer, C.D. (2000), *Matrix Analysis and Applied Linear Algebra*, Philadelphia, PA: American Mathematical Society.

Mueller-Frank, M. and C. Neri (2013), "Social Learning in networks: Theory and experiments," Unpublished manuscript, University of Navarra.

Mouw, T., Chavez, S., Edelblute, H. and A. Verdery (2014), "Binational social networks and assimilation: A test of the importance of transnationalism," *Social Problems* 61, 329–359.

Myles, G.D. and R.A. Naylor (1996), "A model of tax evasion with group conformity and social customs," *European Journal of Political Economy* 12, 49?-66.

Ogbu, J.U. (1997), "Beyond language: Ebonics, proper English and identity in a black American speech community," Unpublished manuscript, University of California at Berkeley, Department of Anthropology.

Panebianco, F. (2014), "Socialization networks and the transmission of interethnic attitudes," *Journal of Economic Theory* 150, 583-?610.

Panebianco, F. and T. Verdier (2017), "Paternalism, homophily and cultural transmission in random networks," *Games and Economic Behavior* 105, 155–176.

Patacchini, E. and Y. Zenou (2012), "Juvenile delinquency and conformism," *Journal of Law, Economic, and Organization* 28, 1–31.

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

Posner, E. (2000), "Law and social norms: The case of tax compliance," *Virginia Law Review* 86, 1781–1819.

Prummer, A. and J.-P. Siedlarek (2017), "Community leaders and the preservation of cultural traits," *Journal of Economic Theory* 168, 143–176.

Prummer, A. (2018), "Religious and cultural leaders," In: J.-P. Carvalho, S. Iyer and J. Rubin (Eds.), *Advances in the Economics of Religion*, Palgrave Macmillan, forthcoming.

Salins, P.D. (1997), *Assimilation, American Style*, New York: Harper Collins.

Schalk-Soekar, S.R., van de Vijver, F.J. and M. Hoogsteder (2004), "Attitudes toward multiculturalism of immigrants and majority members in the Netherlands," *International Journal of Intercultural Relations* 28, 533-?550.

Seneta, E. (1981), *Non-Negative Matrices and Markov Chains*, Second Edition, New York: Springer.

Strang, G. (2016), *Introduction to Linear Algebra*, Fifth Edition, SIAM.

Tahbaz-Salehi, A. and A. Jadbabaie (2008), "A necessary and sufficient condition for consensus over random networks," *IEEE Transactions on Automatic Control* 53, 791?-795.

Verdier, T. and Y. Zenou (2015), "The role of cultural leaders in the transmission of preferences," *Economics Letters* 136, 158–161.

Verdier, T. and Y. Zenou (2017), "The role of social networks in cultural assimilation," *Journal of Urban Economics* 97, 15–39.

Verdier, T. and Y. Zenou (2018), "Cultural leaders and the dynamics of assimilation," *Journal of Economic Theory* 175, 374–414.

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

Appendix: Proofs of the results in the main text

Proof of Proposition 1:

First step: Let us show first that $\mathbf{GM} = \mathbf{MG}$. To do that, we first prove the following lemma

Lemma 1 *If two matrices* $\mathbf{A}$ *and* $\mathbf{B}$ *commute and* $\mathbf{B}$ *is nonsingular, i.e.* $\mathbf{AB} = \mathbf{BA}$*, then* $\mathbf{AB}^{-1} = \mathbf{B}^{-1}\mathbf{A}$.

Proof: We have $\mathbf{AB} = \mathbf{BA}$. This implies that $\mathbf{B}^{-1}\mathbf{AB} = \mathbf{B^{-1}BA}$, which is equivalent to:

$$\mathbf{B}^{-1}\mathbf{AB} = \mathbf{A}$$

This implies that

$$\mathbf{B}^{-1}\mathbf{ABB^{-1}} = \mathbf{AB}^{-1}$$

which is equivalent to

$$\mathbf{B}^{-1}\mathbf{A} = \mathbf{AB}^{-1}$$

This proves the lemma. ∎

Let us now show that $\mathbf{M}^{-1}$ and $\mathbf{G}$ commute, i.e. $\mathbf{M}^{-1}\mathbf{G} = \mathbf{GM}^{-1}$. We have $\mathbf{M} = (\mathbf{I} - \theta\,\mathbf{G})^{-1}$. Thus, $\mathbf{M}^{-1} = (\mathbf{I} - \theta\,\mathbf{G})$. As a result,

$$\mathbf{M}^{-1}\mathbf{G} = (\mathbf{I} - \theta\,\mathbf{G})\,\mathbf{G} = \left(\mathbf{G} - \theta\,\mathbf{G}^2\right) = \mathbf{G}\,(\mathbf{I} - \theta\,\mathbf{G}) = \mathbf{GM}^{-1}$$

Denote $\mathbf{A} = \mathbf{G}$ and $\mathbf{B} = \mathbf{M}^{-1}$. Since $\mathbf{G}$ and $\mathbf{M}^{-1}$ commute, then Lemma 1 shows that $\mathbf{GM} = \mathbf{MG}$.

Second step: Let us show that $\mathbf{TG} = \mathbf{GT}.$ We have that

$$\mathbf{T} = \frac{\gamma}{(1+w)}\mathbf{M} + (1-\gamma)\mathbf{I}$$

This implies that

$$\begin{aligned}
\mathbf{TG} &= \left[\frac{\gamma}{(1+w)}\mathbf{M} + (1-\gamma)\mathbf{I}\right]\mathbf{G} \\
&= \frac{\gamma}{(1+w)}\mathbf{MG} + (1-\gamma)\mathbf{G}
\end{aligned}$$

Since $\mathbf{MG} = \mathbf{GM}$, this can be written as:

$$\begin{aligned}\mathbf{TG} &= \frac{\gamma}{(1+w)}\mathbf{GM} + (1-\gamma)\mathbf{G} \\ &= \mathbf{G}\left[\frac{\gamma}{(1+w)}\mathbf{M} + (1-\gamma)\mathbf{I}\right] \\ &= \mathbf{GT}\end{aligned}$$

Third step: Let us show that $\lim_{t\to\infty} \mathbf{T}^t = \lim_{t\to\infty} \mathbf{G}^t$. Assume that $\mathbf{G}$ is diagonalizable and aperiodic.[22] This implies that $\mathbf{T}$ is diagonalizable. Indeed, if $\mathbf{G}$ is diagonalizable, $\mathbf{M}$ is diagonalizable being a polynomial of $\mathbf{G}$. Since $\mathbf{T}$ is a linear convex combination of $\mathbf{I}$ and $\mathbf{M}$, two diagonalizable matrices, it is also diagonalizable.

Then, since we have seen that $\mathbf{G}$ and $\mathbf{T}$ commute, then they have the same eigenvectors (this is a standard result in linear algebra; see e.g. Strang (2016)). We know that $\mathbf{G}$ converges because $\mathbf{G}$ is aperiodic. Then, $\mathbf{G}$ and $\mathbf{T}$ have the same eigenvector associated with the maximum eigenvalue (which is 1 here), which we denote by $\mathbf{e}$. Since both $\mathbf{T}$ and $\mathbf{G}$ are row-normalized, they both have the same largest eigenvalue equal to 1. This implies that:

$$\mathbf{e}^{\mathrm{T}}\,\mathbf{T} = \mathbf{e}^{\mathrm{T}}\,\mathbf{G} = \mathbf{1}\,\mathbf{e}^{\mathrm{T}}$$

This proves part (i) of the proposition.

Assume now that $\mathbf{G}$ is periodic. Observe that $\mathbf{G}$ and $\mathbf{G}_\epsilon$ commute. Indeed

$$\mathbf{G}\mathbf{G}_\epsilon = \mathbf{G}[\epsilon\mathbf{I} + (1-\epsilon)\mathbf{G}] = \epsilon\mathbf{G} + (1-\epsilon)\mathbf{G}^2 = [\epsilon\mathbf{I} + (1-\epsilon)\mathbf{G}]\mathbf{G} = \mathbf{G}_\epsilon\mathbf{G}$$

Then, by substituting $\mathbf{G}_\epsilon$ to $\mathbf{G}$ and using exactly the same proof as for the case when $\mathbf{G}$ was aperiodic, we obtain the proof of part (ii) of the proposition. ∎

Proof of Proposition 2: Consider, first, the dynamics of norms in (4). It is straightforward to see that, in steady state, $\mathbf{s}^\infty = \mathbf{x}^\infty$ when $\mathbf{s^{t+1}} = \mathbf{s^t} = \mathbf{s}^\infty$ and $\mathbf{x^{t+1}} = \mathbf{x^t} = \mathbf{x}^\infty$. Consider equation (3) that we report here:

$$\mathbf{x}^t = \left(\frac{1}{1+w}\right)[\mathbf{I} - \theta\,\mathbf{G}]^{-1}\,\mathbf{s}^t \tag{40}$$

Then, in steady state,

[22] With some abuse of notation, we say that $\mathbf{G}$ is aperiodic if the submatrix associated to each closed communication class is aperiodic.

$$\mathbf{x}^{\infty} = \left(\frac{1}{1+w}\right)[\mathbf{I} - \theta\,\mathbf{G}]^{-1}\,\mathbf{x}^{\infty} \tag{41}$$

We can write is as:

$$(1+w)\mathbf{x}^{\infty} + \omega\mathbf{G}\mathbf{x}^{\infty} = \mathbf{x}^{\infty} \tag{42}$$

With some algebra, it is immediate to have $\mathbf{x}^{\infty} = \mathbf{G}\mathbf{x}^{\infty}$. This means that $\mathbf{x}^{\infty}$ is the right eigenvector associated to the unit eigenvalue since $\mathbf{G}$ is row-stochastic.
Consider now the utility function in steady state. Since $\mathbf{x}^{\infty} = \mathbf{G}\mathbf{x}^{\infty}$, then $(x_i^{\infty} - \sum_j g_{ij}x_j^{\infty})^2$. Moreover, in steady state $x_i^{\infty} = s_i^{\infty}$. Then the utility is null for all agents, and the welfare is maximal. ▮

Proof of Proposition 3: From (2), we have:

$$x_i^t = \left(\frac{1}{1+\omega}\right) s_i^t + \left(\frac{\omega}{1+\omega}\right) \sum_j g_{ij}x_j^t$$

If we substitute this value into the utility function (1), we obtain:

$$\begin{aligned}
u_i^t &= -\omega\Big[\left(\frac{1}{1+\omega}\right) s_i^t + \left(\frac{\omega}{1+\omega}\right)\sum_j g_{ij}x_j^t - \sum_j g_{ij}x_j^t\Big]^2 \\
&\quad -\Big[\left(\frac{1}{1+\omega}\right) s_i^t + \left(\frac{\omega}{1+\omega}\right)\sum_j g_{ij}x_j^t - s_i^t\Big]^2 \\
&= -\omega\Big[\left(\frac{1}{1+\omega}\right) s_i^t - \left(\frac{1}{1+\omega}\right)\sum_j g_{ij}x_j^t\Big]^2 - \Big[\left(-\frac{\omega}{1+\omega}\right) s_i^t + \left(\frac{\omega}{1+\omega}\right)\sum_j g_{ij}x_j^t\Big]^2 \\
&= \frac{\omega}{(1+\omega)^2}\Big[s_i^t - \sum_j g_{ij}x_j^t\Big]^2 - \frac{\omega^2}{(1+\omega)^2}\Big[s_i^t - \sum_j g_{ij}x_j^t\Big]^2 \\
&= -\frac{\omega}{1+\omega}\Big[s_i^t - \sum_j g_{ij}x_j^t\Big]^2
\end{aligned}$$

This is minimized if either $\omega = 0$ or $\mathbf{s}^t = \mathbf{G}\mathbf{x}^t$. Recall that, in equilibrium, using (3), we have:

$$\mathbf{x}^{t} = \left(\frac{1}{1+\omega}\right)[\mathbf{I} - \theta\,\mathbf{G}]^{-1}\,\mathbf{s}^{t}$$

Therefore, to maximize welfare, we should have:

$$\mathbf{s}^t = \left(\frac{1}{1+\omega}\right) \mathbf{G} \left[\mathbf{I} - \theta\, \mathbf{G}\right]^{-1} \mathbf{s}^t = \mathbf{A}\mathbf{s}^t$$

This means that $\mathbf{s}^t = \mathbf{A}\mathbf{s}^t$, that is $\mathbf{s}^t$ is the right eigenvector associated to $\mathbf{A}$, i.e. $\mathbf{s}^t = \mathbf{e}(\mathbf{A})$. ▮

Proof of Proposition 4: We assume that links in $\bar{\mathbf{G}}$ are $\bar{g}_{ij} \in \{0,1\}$ and that the matrix $\mathbf{G}$ is the row normalization of $\bar{\mathbf{G}}$. Call $i = GL$ the government leader and $j = CL$ the cultural leader. Then, from the characterization results (see Proposition 9 in the Online Appendix B.4), we have:

$$\mathbf{q}_i^{\infty} = \left[\mathbf{I} - (\mathbf{Q}_{CL} + ik)\right]^{-1} \mathbf{q}_i$$

where the matrix $(\mathbf{Q}_{CL} + ik)$ is the bottom right block of the canonical form of the network in which we consider the new link to the native. Then, since there are just two stubborn agents being the only two closed communication classes,

$$\mathbf{q}_j^{\infty} = \left[\mathbf{I} - (\mathbf{Q}_{CL} + ik)\right]^{-1} \mathbf{q}_j = \mathbf{1} - \mathbf{q}_i^{\infty}$$

If we want to maximize the effect of the new link we need to look at the aggregate weight that the stubborn government leader has on the long run distribution. In other words, we want to maximize the sum of the entries of $\mathbf{q}_i^{\infty}$. Notice that since i has only one link to agent k, then

$$\mathbf{q}_i^{'} = \left[0, \ldots, 0, \frac{1}{d_k}, 0, \ldots, 0\right]$$

where d_k is the degree in the network $\mathbf{G}_{CL} + ik$. Then, to maximize the aggregate effect, we choose to link i to the k that maximizes the following

$$\frac{1}{d_k} \sum_{w=1}^{n} m_{wk} \left(\mathbf{G}_{CL} + ik\right)$$

This completes the proof. ▮

Proof of Proposition 5: Recalling Proposition 9, we can write the Markov process associated with $\tilde{\mathbf{G}}$ by reconsidering equation (B.5) for our case. This turns out to be written as

$$\hat{\mathbf{s}}^{\infty} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ \mathbf{g}^{\infty}_{\boldsymbol{f_1}} & \mathbf{g}^{\infty}_{\boldsymbol{f_2}} & \cdots & \mathbf{g}^{\infty}_{\boldsymbol{f_n}} & \mathbf{O} \end{pmatrix} \hat{\mathbf{s}}^{(0)} \tag{43}$$

where $\mathbf{g}^{\infty}_{\boldsymbol{f_i}}$ is the vector of weights assigned by each agent in N to the specific fictitious agent f_i. Let us now consider how each of these vectors looks like. Using the same results as in Proposition 9, we have:

$$\mathbf{g}^{\infty}_{f_i} = [\mathbf{I} - \theta\mathbf{G}]^{-1}\mathbf{g}_{f_i} \tag{44}$$

where $\mathbf{g}_{f_i}$ is the column vector of weights that agents in N assigns to the fictitious f_i. We now recall that only agent i assigns a positive weight to f_i and that this is equal to $\frac{1}{1+w}$. Then, by denoting by m_{ij} the generic entry of matrix $[\mathbf{I} - \theta\mathbf{G}]^{-1}$, and by $\mathbf{m}_{*\mathbf{i}}$, its i^{th} column, we obtain:

$$\mathbf{g}^{\infty}_{f_i} = [\mathbf{I} - \theta\mathbf{G}]^{-1}\mathbf{g}_{f_i} = \frac{1}{1+w}\mathbf{m}_{*\mathbf{i}} \tag{45}$$

We then can derive the following:

$$[\mathbf{g}^{\infty}_{1_\alpha}|\mathbf{g}^{\infty}_{2_\alpha}|\ldots|\mathbf{g}^{\infty}_{n_\alpha}] = \frac{1}{1+w}[\mathbf{I} - \theta\mathbf{G}]^{-1} \tag{46}$$

which is row-normalized. This proves the first part of the proposition.

To find the equilibrium actions just plug $\mathbf{s}^{\infty}$ into (20). This completes the proof. ∎

Proof of Proposition 6: Consider equation (28). By simple algebra, we obtain:

$$\boldsymbol{\alpha} = (1+w)[\mathbf{I} - \theta\mathbf{G}]\mathbf{s}^{\infty} \tag{47}$$

Since $\theta = w/(1+w)$, we have:

$$\boldsymbol{\alpha} = [(1+w)\mathbf{I} - w\mathbf{G}]\mathbf{s}^{\infty} = \mathbf{s}^{\infty} + w(\mathbf{I} - \mathbf{G})\mathbf{s}^{\infty} \tag{48}$$

Given a target $\mathbf{s}^*$, we immediately have that the optimal $\boldsymbol{\alpha^*}$ is $\boldsymbol{\alpha^*} = \mathbf{s}^* + w(\mathbf{I} - \mathbf{G})\mathbf{s}^*$. Then

$$\boldsymbol{\sigma^*} = \mathbf{s}^* - \boldsymbol{\alpha} + w(\mathbf{I} - \mathbf{G})\mathbf{s}^* \tag{49}$$

Notice however that, if for all $i, j \in I$, $s_i^* = s_j^*$, then $(\mathbf{I} - \mathbf{G})\mathbf{s}^* = 0$ since every agent has a target norm equal to the average of her neighbors' targets. Then $\boldsymbol{\sigma^*} = \boldsymbol{s^*} - \boldsymbol{\alpha}$. This proves the result. ∎

Online Appendix
Conformism, Social Norms and the Dynamics of Assimilation

By Gonzalo Olcina[1], Fabrizio Panebianco[2] and Yves Zenou[3]

March 29, 2018

A Some Results in Linear Algebra

We state here (without proving them) some standard results in linear algebra.

Definition 2 (Irreducible matrix) $\mathbf{G}$ *is said to be irreducible if all agents form a unique communication class. Otherwise* $\mathbf{G}$ *is reducible. In terms of network,* $\mathbf{G}$ *is irreducible if it is a strongly connected network .*

LEMMA **2** *Consider* $\mathbf{G}$ *to be an irreducible matrix. Then its largest eigenvalue* $\lambda_1 = 1$ *and there is a positive eigenvector associated to it.*

Definition 3 (Periodicity) *A state in a Markov chain is periodic if the chain can return to the state only at multiples of some integer larger than* 1*. Formally, the period* $d(x)$ *of state* $x \in S$ *is:*

$$d(x) = gcd\{n \in \mathbb{N}_+ : P^n(x,x) > 0\} \tag{A.1}$$

where S *is the state space,* P *the transition probability matrix and gcd stands for greater common divisor. State* x *is aperiodic if* $d(x) = 1$ *and periodic if* $d(x) > 1$

LEMMA **3** *All agents belonging to the same communication class have the same period.*
If a matrix is periodic then $\lim_{t\to\infty} \mathbf{G}^t$ *does not exist and there are precisely* d *complex eigenvectors of length one.*

Definition 4 (Primitive Matrix) *If* $\mathbf{G}$ *is aperiodic and irreducible, it is called* primitive

Definition 5 *Consider the eigenvector associated to the unit eigenvalue of a stochastic matrix* $\mathbf{G}$*. We refer to it as the Perron-Froebenius eigenvector, and we call it* $\mathbf{e}(\mathbf{G})$*.*

[1]Universitat de València and ERI-CES, Spain. Email: Gonzalo.Olcina@uv.es.
[2]Department of Decision Sciences and IGIER, Bocconi University, Italy. E-mail: fabrizio.panebianco@unibocconi.it.
[3]Monash University and CEPR. Email: yves.zenou@monash.edu.

Consider $\mathbf{Q}$ defined in equation (B.2).

LEMMA 4 *Consider a matrix* $\mathbf{G}$ *in its canonical form. Then,* $\lim_{t\to\infty} \mathbf{Q}^t = 0$ *entry wise. This convergence happens geometrically fast.*

B Convergence and Network Topology

B.1 Basic definitions

Let us provide some results of Markov processes that we write in terms of our model, where we do not have transition matrices that determine the probability of switching from one state to another, but, instead, we have the influence of the network.

We now propose some useful definitions.

Definition 6 (Influenced Agents) *Agent* i *is influenced by agent* j *if there exists a sequence of nodes (agents) such that* $g_{ik}g_{kw}\cdots g_{zj} > 0$. *We denote this as* $i \to j$. *If* j *is also influenced by* i *then* $i \leftrightarrow j$.

Notice that $i \to j$ does not mean that g_{ij} are linked, but that there is a path from i to j. Considering, for example, the first network $\mathbf{G_1}$ in Figure 3. Then, we can see that $8 \to 6$ since $g_{87}g_{71}g_{15}g_{56} > 0$.

Definition 7 (Communication Class) *A set* $C \subset N$ *is called a communication class if: (a)* $i, j \in C \Rightarrow i \leftrightarrow j$*; (b)* $i \in C$ *and* $i \leftrightarrow j \Rightarrow j \in C$.

In words, a communication class is a subset of agents who are all (directly or indirectly) influenced by each other. In the networks in Figure 3, the agents belonging to the same communication class have the same color. This means that the first network is composed of agents that all belong to the same communication class while, in the two other networks, each has three different communication classes.

Definition 8 (Closed Communication Class) *A set* $C \subset N$ *is a closed communication class if:* $i \in C$ *and* $j \notin C \Rightarrow i$ *is not influenced by* j.

Given that $g_{ii} = 0$, closed communication classes cannot be formed by singletons. Consider the networks given in Figure 3. In network 2, there is one closed communication class composed of agents 1, 2 and 3 while, in network 3, there are two closed communication classes composed of agents 1, 2, 3 and 4, 5, 6, respectively. In these networks, closed communication classes are represented by circles while other communication classes (so called transient communication classes) by diamonds. Observe that the definition of closed communication classes implies that there are no link between agents belonging to two closed communication classes.

B.2 Convergence for strongly connected networks

B.2.1 General results

We first analyze the case of a *strongly connected* network $\mathbf{G}$. Recall that a network is strongly connected if all nodes belong to a unique closed communication class. This case is relatively standard and has been analyzed, for example, by Golub and Jackson (2010, 2012). There are, however, two main differences with the standard case. First, in the standard case, the steady-state norms are given by the network matrix $\mathbf{G}$. In our case, we have seen that the dynamics of norms is governed by $\mathbf{T}$ instead of $\mathbf{G}$. Second, the standard results only hold for aperiodic $\mathbf{G}$ while we will prove convergence of norms also for periodic $\mathbf{G}$.

PROPOSITION **7** *Assume that the network is strongly connected. If* $\mathbf{G}$ *is aperiodic, then,*

$$\lim_{t\to+\infty} \mathbf{T}^t\,\mathbf{s} = \begin{pmatrix} \mathbf{e}(\mathbf{G}) \\ \cdots \\ \cdots \\ \mathbf{e}(\mathbf{G}) \end{pmatrix} \mathbf{s} \qquad \textit{for every } \mathbf{s} \in [0,1]^n$$

If $\mathbf{G}$ *is periodic, the same result holds by considering the matrix* $\mathbf{G}_\epsilon$ *instead of* $\mathbf{G}$.

The convergence results comes from Markov chain theory together with our finding that the asymptotic properties of $\mathbf{T}$ and $\mathbf{G}$ are the same. While it is well known that, for a strongly connected $\mathbf{G}$, its limit is given by its Perron-Froebenius eigenvector, the way in which the two matrices $\mathbf{G}$ and $\mathbf{T}$ relate to each other is not straightforward, and depends on the way incentives and norms updating process interact. More precisely, the matrix $\mathbf{T}$ is *irreducible* because the associated network $\mathbf{G}$ is assumed to be strongly connected. Moreover, the matrix $\mathbf{T}$ is *aperiodic* because every agent assigns a positive weight to herself in the norm updating process ($\gamma > 0$) and, since $\mathbf{T}$ is a function of the network $\mathbf{G}$, which is strongly connected, there are always self loops. This implies that $t_{ii} > 0$, $\forall i$. Since $\mathbf{T}$ is an irreducible and aperiodic non-negative matrix, it is a *primitive* matrix. This guarantees that the process converges to a unique steady state. Finally, since the matrix $\mathbf{T}$ is *row* *stochastic*, its largest eigenvalue is 1, and therefore, there is a unique left eigenvector $\mathbf{e}(\mathbf{T})$ with positive components such that $\mathbf{e}(\mathbf{T}) = \mathbf{e}(\mathbf{T})\,\mathbf{T}$. The eigenvector property is just saying that $e_i = \sum_{j\in N} t_{ji}e_j$ for all i, so that the opinions of agents with greater influence have a greater weight in the final convergence belief.

Proposition 7 also derives a similar result when the matrix $\mathbf{G}$ is periodic. When this is the case, it is impossible to find $\lim_{t\to\infty}\mathbf{G}^t$. However, for the very same reasons described above, the matrix $\mathbf{T}$ is primitive. Then it is enough to study the perturbed matrix $\mathbf{G}_\epsilon$ instead of $\mathbf{G}$ to obtain the convergence result.

B.2.2 Example

Let us illustrate our convergence result for strongly connected networks stated in Proposition 7. In this proposition, we characterize the unique steady-state norm s_i^∞ for each individual i, which is the same for all individuals in the network, and show that it is equal to a weighted average of the initial norms of all agents, where the weights are each agent's eigenvector centrality. This means that the more central an individual is in the network, the higher is her weight in the determination of the common steady-state norm. Let us illustrate this result for network 1 in Figure 3, which is a strongly connected network with one communication class (i.e. there is no agent outside the communication class) and understand the implications in terms of assimilation norms.

The adjacency matrix of network 1 in Figure 3 is given by:

$$\mathbf{G_1} = \begin{bmatrix} 0 & 0 & \frac{1}{2} & 0 & \frac{1}{2} & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{3} & 0 & \frac{1}{3} & 0 & 0 & \frac{1}{3} \\ 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 & 0 \\ \frac{1}{3} & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{3} & \frac{1}{3} \\ 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} & 0 \end{bmatrix}$$

Assume the following (transpose) vector of initial norms:

$$[\mathbf{s^{(0)}}]^T = \begin{bmatrix} 0.1 & 0.2 & 0.3 & 0.4 & 0.5 & 0.6 & 0.7 & 0.8 & 0.9 \end{bmatrix} \tag{B.1}$$

We expect the norms to converge to the same value, which is given by the weighted average of the initial norms, where the weights are given by the eigenvector centralities of the agents. Assume that $w = \gamma = 0.5$. We easily obtain:

$$\mathbf{T}_{\mathbf{1}}^{\infty} = \begin{pmatrix} \mathbf{e}(\mathbf{G}) \\ ... \\ ... \\ ... \\ ... \\ ... \\ ... \\ ... \\ \mathbf{e}(\mathbf{G}) \end{pmatrix} = \begin{bmatrix} 0.095 & 0.048 & 0.095 & 0.095 & 0.143 & 0.095 & 0.143 & 0.127 & 0.159 \\ 0.095 & 0.048 & 0.095 & 0.095 & 0.143 & 0.095 & 0.143 & 0.127 & 0.159 \\ 0.095 & 0.048 & 0.095 & 0.095 & 0.143 & 0.095 & 0.143 & 0.127 & 0.159 \\ 0.095 & 0.048 & 0.095 & 0.0952 & 0.143 & 0.095 & 0.143 & 0.127 & 0.159 \\ 0.095 & 0.048 & 0.095 & 0.095 & 0.143 & 0.095 & 0.143 & 0.127 & 0.159 \\ 0.095 & 0.048 & 0.095 & 0.095 & 0.143 & 0.095 & 0.143 & 0.127 & 0.159 \\ 0.095 & 0.048 & 0.095 & 0.095 & 0.143 & 0.095 & 0.143 & 0.127 & 0.159 \\ 0.095 & 0.048 & 0.095 & 0.095 & 0.143 & 0.095 & 0.143 & 0.127 & 0.159 \\ 0.095 & 0.048 & 0.095 & 0.095 & 0.143 & 0.095 & 0.143 & 0.127 & 0.159 \end{bmatrix}$$

We can see that the most influencial individual is 9 because she is the more central agent (in terms of eigenvector centrality) in the network. The limit social norm in terms of assimilation will then be the weighted average of all initial norms, where the weights are given by each row of this matrix. It is easily verified that it is given by: $s_i^{\infty} = 0.559$ for all $i = 1, ..., 9$. In other words, even though each ethnic minority has a very different initial assimilation norm, they all end up with the same social norm equal to 0.559, which is slightly favorable to assimilation.

Let us now look at the dynamics of social norms and how they converge in steady state. For that, we examine the convergence of the assimilation norm of individuals 2, 4, and 9, who belong to the same communication class but to different groups and have very different initial assimilation norms. Figure B.1 illustrates this dynamics.

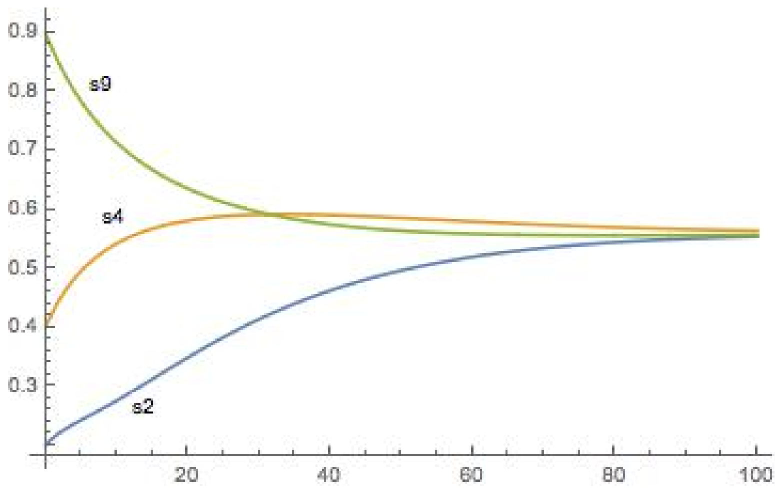

Figure B.1: Norm convergence in a strongly connected network

We can see that these individuals start with very different initial norms (at time $t = 0$) and then, as time passes by, they converge to the same weighted average norm $s^{\infty} = 0.559$. Individual 9 starts with a very high assimilation norm ($s_9^{(0)} = 0.9$) while individual 2 starts with $s_2^{(0)} = 0.2$, a norm not in favor of assimilation (for example, because her parents want to keep their original culture and language and feel threatened by the majority culture norm). Still, because of peer effects, social interactions and the structure of the network, these two agents end up with the same steady-state norm equal to 0.559, which implies a relatively favorable level of assimilation norm.

B.3 Convergence for networks with only one closed communication class

B.3.1 General results

Denote by $\mathbf{C}$ the $k \times k$ (transition) matrix associated to the closed communication class with k agents. Then, for this subset of agents only, the previous result about strongly connected networks holds. We need to understand, however, the convergence norms of all the other agents in the network who do not belong to the closed communication class. Since their norms depend directly or indirectly on that of the agents in the closed communication class, they also need to converge to the norm of the agents in the closed communication class. Denote by $\tilde{\mathbf{e}}(\mathbf{C}) = [\mathbf{e}(\mathbf{C})|0, \cdots, 0]$ the $1 \times n$ vector composed of $\mathbf{e}(\mathbf{C})$ augmented by $n - k$ entries equal to 0.

Proposition 8 *Assume that* $\mathbf{G}$ *only has one closed communication class of* k *agents. Then, in steady state, the norm of all agents in the network will converge to the same value. If* $\mathbf{C}$ *is aperiodic, then*

$$\mathbf{s}^{\infty} = \begin{pmatrix} \tilde{\mathbf{e}}(\mathbf{C}) \\ \cdots \\ \cdots \\ \tilde{\mathbf{e}}(\mathbf{C}) \end{pmatrix} \mathbf{s}^{(0)}$$

If $\mathbf{C}$ *is periodic, then the same result holds by considering the matrix* $\mathbf{C}_{\epsilon}$ *instead of* $\mathbf{C}$.

As above, this result relies on the fact that $\mathbf{G}$ and $\mathbf{T}$ have the same asymptotic properties and that there exists only one closed communication class. Observe that, since the communication class is composed of more than one agent, then the convergence value is given by the Perron-Froebenius eigenvector centrality in the sub-matrix representing the communication class.

B.3.2 Example

Let us now study networks with only one closed communication class. In Proposition 8, we characterize the unique steady-state norm s_i^{∞} of each individual i belonging to the closed communication class and shows that it is equal to a weighted average of the initial norms of the individuals belonging to the closed communication class, where the weights are determined by each agent's eigenvector centrality. Importantly, all individuals who do not belong to the closed communication class will be assigned a weight of zero when determining the common steady-state norm. In other words, in steady-state, all individuals will adhere to the same common assimilation norm, which is calculated by taking the weighted average of all initial norms where the weights are equal to the eigenvector centrality of each individual who belongs to the closed communication class and zero for all the other individuals who do *not* belong to the closed communication class.

Consider network 2 in Figure 3, which has only one closed communication class composed of agents 1, 2, 3. Thus, we expect the norm of all agents outside of the closed communication class to converge to the common steady-state assimilation norm of agents 1, 2, 3. The adjacency matrix of network 2 is given by:

$$
\mathbf{G_2} = \begin{bmatrix}
0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
\frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 \\
0 & 0 & 0 & \frac{1}{3} & 0 & \frac{1}{3} & 0 & 0 & \frac{1}{3} \\
0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 & 0 \\
\frac{1}{3} & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{3} & \frac{1}{3} \\
0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{2} & 0 & \frac{1}{2} \\
0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} & 0
\end{bmatrix}
$$

The matrix of the closed communication class $C = \{1, 2, 3\}$ is given by:

$$
\mathbf{C} = \begin{bmatrix}
0 & 0 & 1 \\
0 & 0 & 1 \\
\frac{1}{2} & \frac{1}{2} & 0
\end{bmatrix}
$$

and is periodic (period 2). To show the convergence, we need to determine $\mathbf{C}_\epsilon$ (see equation 9), with, for example, $\epsilon = 0.1$. Then,

$$
\mathbf{C}_\epsilon = \begin{bmatrix}
\frac{1}{10} & 0 & \frac{9}{10} \\
0 & \frac{1}{10} & \frac{9}{10} \\
\frac{9}{20} & \frac{9}{20} & \frac{1}{10}
\end{bmatrix}
$$

From $\mathbf{C}_\epsilon$ we can calculate $\mathbf{G}_\epsilon$. For that, in $\mathbf{G_2}$, we replace $\mathbf{C}$ by $\mathbf{C}_\epsilon$. In that case, the convergence matrix is given by:

$$
\mathbf{T_2^\infty} = \begin{pmatrix}
\tilde{\mathbf{e}}(\mathbf{G}) \\
... \\
... \\
... \\
... \\
... \\
... \\
... \\
\tilde{\mathbf{e}}(\mathbf{G})
\end{pmatrix} = \begin{bmatrix}
0.25 & 0.25 & 0.5 & 0 & 0 & 0 & 0 & 0 & 0 \\
0.25 & 0.25 & 0.5 & 0 & 0 & 0 & 0 & 0 & 0 \\
0.25 & 0.25 & 0.5 & 0 & 0 & 0 & 0 & 0 & 0 \\
0.25 & 0.25 & 0.5 & 0 & 0 & 0 & 0 & 0 & 0 \\
0.25 & 0.25 & 0.5 & 0 & 0 & 0 & 0 & 0 & 0 \\
0.25 & 0.25 & 0.5 & 0 & 0 & 0 & 0 & 0 & 0 \\
0.25 & 0.25 & 0.5 & 0 & 0 & 0 & 0 & 0 & 0 \\
0.25 & 0.25 & 0.5 & 0 & 0 & 0 & 0 & 0 & 0 \\
0.25 & 0.25 & 0.5 & 0 & 0 & 0 & 0 & 0 & 0
\end{bmatrix}
$$

Individual 3, who is the most central agent in the closed communication class, has the highest weight (i.e. 0.5). If we consider the same vector of initial norms as in (B.1), then, the limit assimilation norm will be equal to:

$$s_i^{\infty} = 0.25 s_1^{(0)} + 0.25 s_2^{(0)} + 0.5 s_3^{(0)} = 0.225, \quad \text{for all } i\text{=1,...,9}$$

Clearly, all indidividuals who are not in the closed communication class (i.e. agents 4, ..., 9) have no influence on this steady-state norm but will still adopt it. Indeed, as stated above, if there is only one closed communication class, then this class produces a system whose dynamics is independent of the dynamics of the other individuals in the network not belonging to this closed communication class.

Let us now examine the dynamics of social norms and how they converge in steady state. For that, we look at the convergence of the norm of individuals 2, 4, and 9, who belong to different groups (only 2 belongs to the closed communication class). Figure B.2 displays these dynamics.

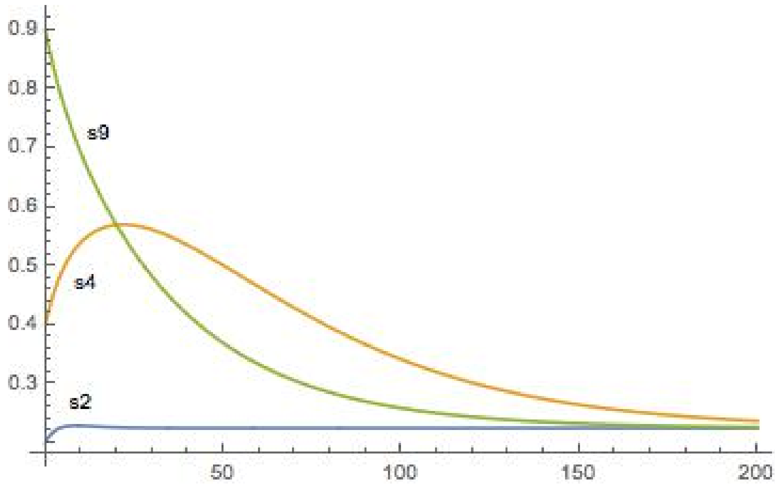

Figure B.2: Norm convergence with one communication class

We can see that individual 4, who is the furthest away from the closed communication class and who starts with an initial assimilation norm of 0.4 follows first the social norm of indidivual 9. Then, eventually, she converges to the norm of the closed communication class. On the contrary, indidivual 9 is not influenced by individual 4 because no agent in 9's communication class is influenced by any agent from 4's communication class. What is interesting here is that, in terms of assimilation norms, agents 4, ..., 9, who have inherited a relatively high assimilation norm (especially agents 7, 8, 9) will end up having a steady-state norm of 0.225, which is not at all favorable to assimilation. This is due to their positions in the network, which are peripheral, and the fact that the agents who matter (those belonging to the closed communication class) have low initial social norms in terms of assimilation.

B.4 Convergence for any network

Let us now state our result on convergence of social norms for any possible network. Given any network $\mathbf{G}$, we can partition the set of agents such that $N = C_1 \cup C_2 \cup \cdots \cup C_m \cup Q$ where $C_1 \cdots C_m$ are closed communication classes while Q is formed by all the remaining agents. Moreover we call C_p a generic communication class, $n_p := |C_p|$ and $n_Q := |Q| = n - \sum_{p=1}^{m} n_p$.

Consider now a network $\mathbf{G}$. By columns and rows transposition, we can write the matrix $\mathbf{G}$ as an upper triangular block matrix, in its *Canonical Form*, as follows:

$$\widetilde{\mathbf{G}} = \begin{pmatrix} \mathbf{C_1} & 0 & 0 & 0 & 0 \\ 0 & \mathbf{C_2} & 0 & 0 & 0 \\ 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & \mathbf{C_m} & 0 \\ \mathbf{Q}_{C_1} & \mathbf{Q}_{C_2} & \mathbf{Q}_{C_{\dots}} & \mathbf{Q}_{C_m} & \mathbf{Q} \end{pmatrix} \tag{B.2}$$

where $\mathbf{C_p}$ is the transition matrix associated to the C_p closed communication class, while the last rows report the matrices representing how much agents not belonging to any C_p are influenced by those agents in a generic C_p, for $i = 1, \dots, m$. Notice that the generic $\mathbf{Q}_{C_p}$ has dimension $n_Q \times n_p$, that every $\mathbf{C}_p$ is stochastic, and that $\mathbf{Q}$ is substochastic.

To illustrate this, consider network 3 in Figure 3. Its adjacency matrix is given by:

$$\mathbf{G_3} = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{2} & 0 & \frac{1}{2} & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 & 0 \\ \frac{1}{3} & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{3} & \frac{1}{3} \\ 0 & 0 & 0 & 0 & 0 & \frac{1}{3} & \frac{1}{3} & 0 & \frac{1}{3} \\ 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} & 0 \end{bmatrix} \tag{B.3}$$

where this matrix keeps track of *outdegrees* only and where the sum of each row is equal to one (row-normalized matrix). Consider, for example, the first row corresponding to agent 1. Since we only consider outdegrees, the link between 1 and 7 does not appear here and it only exists in row 7.

It is easily verified that this network has two closed communication classes, $C_1 = \{1, 2, 3\}$ and $C_2 = \{4, 5, 6\}$ and thus $Q = \{7, 8, 9\}$, which means that agents $7, 8, 9$ do not belong to any closed communication class. Then, the canonical form of $\mathbf{G_3}$ is given by $\widetilde{\mathbf{G}}_3$, where

$$\mathbf{C_1} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 1 \\ \frac{1}{2} & \frac{1}{2} & 0 \end{bmatrix}; \mathbf{C_2} = \begin{bmatrix} 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & 0 \end{bmatrix}; \mathbf{Q} = \begin{bmatrix} 0 & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & 0 & \frac{1}{3} \\ \frac{1}{2} & \frac{1}{2} & 0 \end{bmatrix}; \mathbf{Q}_{C_1} = \begin{bmatrix} \frac{1}{3} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}; \mathbf{Q}_{C_2} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & \frac{1}{3} \\ 0 & 0 & 0 \end{bmatrix}. \tag{B.4}$$

In particular, $\mathbf{Q}_{C_1}$, which keeps track of the links between individuals who do not belong to any closed communication class, i.e. individuals $7, 8, 9$, and those who belong to $C_1 = \{1, 2, 3\}$, shows that there is only one direct link between individuals 7 and 1 with weight 1/3. Similarly, $\mathbf{Q}_{C_2}$, which keeps track of the links between individuals who do not belong to any closed communication class, i.e. individuals $7, 8, 9$, and those who belong to $C_2 = \{4, 5, 6\}$, shows that there is only one direct link between individuals 8 and 6 with weight 1/3.

Definition 9 *Consider the left eigenvector associated to the unit eigenvalue of a stochastic matrix* $\mathbf{G}$*. This is the Perron-Froebenius eigenvector, and we call it* $\mathbf{e}(\mathbf{G})$*, which is defined as:* $\mathbf{e}^T(\mathbf{G}) = \mathbf{e}^T(\mathbf{G})\mathbf{G}$.

Recall that $\mathbf{e}^T(\mathbf{G})$ is a row vector, and, given that $\mathbf{G}$ is a stochastic matrix, has positive entries whose sum is normalized to one. We are now ready to characterize the steady-state norms in terms of the structure of the network $\mathbf{G}$. We first recall some well-known results stating that, if the network is strongly connected, then in steady state norms converge to the same value (see Appendix B.2). This result easily extends to the case of network with just one closed communication class (see Appendix B.3). Consider now the case of a generic network $\mathbf{G}$.

While the convergence level of norms of agents belonging to a closed communication class can be derived directly from previous results on strongly connected networks, the characterization of norms of agents not belonging to any closed communication class is more challenging. Define $\mathbf{O}$ as a square matrix of zeros of dimension n_q. Denote by $\mathbf{e}(\mathbf{C_p})$ the Perron-Frobenius eigenvector for the communication class C_p, and denote by $e_i(\mathbf{C_p})$ its i^{th} entry.

Proposition 9 *Consider any network* $\mathbf{G}$. *Each agent* i *maximizes utility (1) and her norm follows the updating process described in (5). Assume* $\mathbf{G}$ *to be aperiodic. Then, in steady state,*

$$\mathbf{s}^{\infty} = \begin{pmatrix} \mathbf{C_1}^{\infty} & 0 & 0 & 0 & 0 \\ 0 & \mathbf{C_2}^{\infty} & 0 & 0 & 0 \\ 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & \mathbf{C_m}^{\infty} & 0 \\ \mathbf{Q}^{\infty}_{C_1} & \mathbf{Q}^{\infty}_{C_2} & \mathbf{Q}^{\infty}_{C_{\dots}} & \mathbf{Q}^{\infty}_{C_m} & \mathbf{O} \end{pmatrix} \mathbf{s}^{(0)} \tag{B.5}$$

where:

$$\mathbf{C}^{\infty}_{\mathbf{p}} = \begin{pmatrix} \mathbf{e}(\mathbf{C_p}) \\ \cdots \\ \cdots \\ \mathbf{e}(\mathbf{C_p}) \end{pmatrix}^{n_p \times n_p}, \ \forall p = 1, \dots, m. \tag{B.6}$$

$$\mathbf{Q}^{\infty}_{C_p} = [\mathbf{q}^{\infty}_{1_p} | \mathbf{q}^{\infty}_{2_p} | \dots | \mathbf{q}^{\infty}_{n_{p_p}}]^{n_q \times n_p}, \ \forall p = 1, \dots, m. \tag{B.7}$$

and each column is given by the vector

$$\mathbf{q}^{\infty}_{i_p} = e_i(\mathbf{C_p})[\mathbf{I} - \mathbf{Q}]^{-1} \cdot \mathbf{Q}_{C_p} \cdot \mathbf{1}, \ \forall i \in C_p, \ \forall p = 1, \dots, m \tag{B.8}$$

If any $\mathbf{C_p}$ *is periodic, then the same result holds by considering the corresponding perturbed matrix* $\mathbf{C_{p\epsilon}}$.

This result shows that the norms of agents not belonging to any closed communication class converge to some average of the others' norms, which is given by (B.5). Notably this result holds independently of $\mathbf{G}$ being periodic or not. Let us now show how the limit matrix is calculated.

Consider, first, the agents belonging to any of the closed communication classes C_p, identified by the matrix $\mathbf{C_p}$ of dimension n_p. Then, the long-run norms of agents belonging to C_p are given by $\mathbf{e}(\mathbf{C_p})$. Thus, we can characterize the block-diagonal matrix $\mathbf{C}^{\infty}_{\mathbf{p}}$ as in (B.6).

Next, consider the agents who do not belong to any closed communication class, characterized by the matrix $\mathbf{Q}$. Notice that $\mathbf{Q}$ characterizes the network existing only among these agents, disregarding the links from and towards any other agent belonging to any closed communication class. In other words, $\mathbf{Q}$ is the adjacency matrix of the network $\mathbf{G}$ restricted to agents in Q (i.e. agents not belonging to any closed communication class). Consider now agents $h \in Q$, and $i \in C_p$. We would like to determine the weight that h assigns to i. In (B.7), $\mathbf{Q}_{C_p}^{\infty}$ gives the matrix of weights in steady state that all agents in Q assign to agents in C_p. The ith column of this matrix, denoted by $\mathbf{q}_{i_p}^{\infty}$ and reported in (B.8), represents the weigths that agents in Q assign to agent $i \in C_p$. Their steady-state norm value depends on the weights directly or indirectly assigned to other agents in $\mathbf{Q}$ times how much each of them is directly linked to agents in any C_p times $e_i(\mathbf{C_p})$, that is how much $i \in C_p$ weights in the final norms of agents in C_p.

Proposition 9 describes this result. Indeed, consider equation (B.8) and focus on how $\mathbf{q}_{i_p}^{\infty}$ is constructed. The matrix $[\mathbf{I} - \mathbf{Q}]^{-1}$, which is convergent given that $\mathbf{Q}$ is a substochastic matrix, is a Neumann series representing how much any two agents in Q directly or indirectly assign weight to each other. Now, select another agent $k \in \mathbf{Q}$ and call m_{hk} the generic entry of $[\mathbf{I} - \mathbf{Q}]^{-1}$. Then, m_{hk} gives how many direct or indirect walks h has towards k. Then, fixing agent $i \in C_p$, we have:

$$e_i(\mathbf{C_p})[\mathbf{I} - \mathbf{Q}]^{-1}\mathbf{Q}_{C_p}\mathbf{1} = e_i(\mathbf{C_p}) \sum_{k \in Q} m_{hk} \sum_{j \in C_p} g_{kj}, \ \forall h \in Q \ .$$

The weight $h \in Q$ assigns to $i \in C_p$ depends on how many paths $h \in Q$ has towards any other $k \in Q$, times how much weight any $k \in Q$ assigns to agents in C_p times how much weight $i \in C_p$ has in the final norm of C_p. $[\mathbf{I} - \mathbf{Q}]^{-1}$ is usually called the *Fundamental Matrix of Absorbing Chains*. By doing so, in terms of our model, we can fully characterize the steady-state level of norms for any network.

Let us summarize how we calculate the steady-state norms in any network. Consider a network where there are m closed communication classes, where, in each closed communication classes C_p, there are n_p individuals, and n_Q individuals who do not belong to any closed communication class. If the total number of individuals in the network is n, then $n = \sum_{p=1}^{p=m} n_p + n_Q$. First, we need to write the adjacency matrix $\mathbf{G}$ in its canonical form $\tilde{\mathbf{G}}$ as in equation (B.2). In $\tilde{\mathbf{G}}$, there are three types of matrices that keep track of the weighted links between all agents in the network. First, there are the squared matrices $\mathbf{C}_p$, for $p = 1, ..., m$ (each has a dimension $n_p \times n_p$) corresponding to the closed communication classes. Second, there is the squared matrix $\mathbf{Q}$ (which is of dimension $n_Q \times n_Q$) of all agents who do not belong to any closed communication class. Third, there are the matrices $\mathbf{Q}_{\mathbf{C}_p}$, for $p = 1, ..., m$ (each has a dimension $n_Q \times n_p$), which keep track of the links between individuals who do not belong to any closed communication class and those who belong to the communication class C_p.

Given the matrix $\widetilde{\mathbf{G}}$, Proposition 9 characterizes the social norms of all individuals in steady state. First, for all individuals belonging to the same closed communication class C_p, they all converge to the same steady state norm, which is a weighted average of their initial beliefs and where each weight is the eigenvector centrality of each individuals belonging to C_p. Second, for all the other individuals who do not belong to any closed communication class, their limiting norms are calculated as in (B.7) and (B.8), which mean that the norms of these individuals converge to some average of the norms of agents belonging to the closed communication classes. However, this convergence depends on their position in the network and if there is a path between a given individual not belonging to any closed communication class and someone else in a closed communication class.

B.4.1 Convergence for any network: An example

Let us provide an example to illustrate the results stated in Proposition 9.[4] In this proposition, we show that different norms will emerge in steady state. Indeed, if there are m closed communication classes and n_Q individuals who do not belong to any closed communication class, then there will be $m + n_Q$ steady-state norms.

Consider network 3 in Figure 3. Its adjacency matrix is given by (B.3). Let us now determine the steady state norms of all individuals in the network. As stated above, since there are two closed communication classes and three individuals who do not belong to any closed communication class, there will be five social norms in steady state. It is easily verified that the communication class C_1 is periodic (period 2) so that we will use $C_{1\epsilon}$ with $\epsilon = 0.10$. Using (9), we have:

$$\mathbf{C_{1\epsilon}} = \begin{bmatrix} 0.1 & 0 & 0.9 \\ 0 & 0.1 & 0.9 \\ 0.45 & 0.45 & 0.1 \end{bmatrix}$$

We obtain:

[4]In appendices B.2 and B.3, we provide examples where we calculate the convergence of social norms for strongly connected networks and networks with one closed communicaton class, respectively.

$$
\mathbf{T}_{\mathbf{3}}^{\infty} = \begin{bmatrix}
0.25 & 0.25 & 0.5 & 0 & 0 & 0 & 0 & 0 & 0 \\
0.25 & 0.25 & 0.5 & 0 & 0 & 0 & 0 & 0 & 0 \\
0.25 & 0.25 & 0.5 & 0 & 0 & 0 & 0 & 0 & 0 \\
0. & 0. & 0. & 0.333 & 0.333 & 0.333 & 0 & 0 & 0 \\
0. & 0. & 0. & 0.333 & 0.333 & 0.333 & 0 & 0 & 0 \\
0. & 0. & 0. & 0.333 & 0.333 & 0.333 & 0 & 0 & 0 \\
0.156 & 0.156 & 0.313 & 0.125 & 0.125 & 0.125 & 0 & 0 & 0 \\
0.0938 & 0.0938 & 0.188 & 0.208 & 0.208 & 0.208 & 0 & 0 & 0 \\
0.125 & 0.125 & 0.25 & 0.167 & 0.167 & 0.167 & 0 & 0 & 0
\end{bmatrix} \tag{B.9}
$$

We see that the limit norm of the closed communication class $C_1 = \{1, 2, 3\}$ will only take into account the weights of individuals $1, 2, 3$ while that of the closed communication class $C_2 = \{4, 5, 6\}$ will only take into account the weights of individuals $4, 5, 6$. We also see that, to calculate the limit norm of each individual who do not belong to any closed communication class, i.e. individuals $7, 8, 9$, one puts a weight of zero for each of them and puts a positive weight for each other individual belonging to either C_1 or C_2. Interestingly, the weights depend on the position in the network and the distance between the individual who does not belong to any closed communication class and those who belong to a closed communication class. Take, for example, individual 7 (row 7 in $\mathbf{T}_{\mathbf{3}}^{\infty}$). In the determination of her limit norm, the highest weight has been put on individual 3 because there is a direct link between individual 7 and someone from C_1 (i.e. individual 1), and individual 3 has the highest eigenvector centrality (0.5) of all individuals in C_1. The lowest weight has been put on individual 4, 5 or 6 because there is no direct link between individual 7 and anybody from C_2 and all individuals in C_2 have the same eigenvector centrality equal to 0.333.

To better understand this result, let us see how the matrix $\mathbf{T}_{\mathbf{3}}^{\infty}$ is constructed. First, we calculate the matrix that keeps track of the walks of different lengths in the network only for individuals who do not belong to any closed communication class (i.e. individuals $7, 8, 9$):

$$
[\mathbf{I} - \mathbf{Q}]^{-1} = \begin{bmatrix}
\frac{15}{8} & \frac{9}{8} & 1 \\
\frac{9}{8} & \frac{15}{8} & \frac{8}{3} \\
\frac{3}{2} & \frac{3}{2} & 2
\end{bmatrix}
$$

Consider, for example, the weight each agent in Q assigns to agent 1 (who belongs to C_1) in steady state. This is given by $\mathbf{q}_{i_1}^{\infty}$, for $i = 7, 8, 9$. In Proposition 9, we show that: $\mathbf{q}_{i_1}^{\infty} = e_1(\mathbf{C_1})[\mathbf{I} - \mathbf{Q}]^{-1} \cdot \mathbf{Q}_{C_1} \cdot \mathbf{1}$. We have: $e_1(\mathbf{C_1}) = 0.25$. Then, we easily obtain:

$$\mathbf{q}_{i_1}^{\infty} = 0.25 \begin{bmatrix} \frac{15}{8} & \frac{9}{8} & 1 \\ \frac{9}{8} & \frac{15}{8} & \frac{8}{3} \\ \frac{3}{2} & \frac{3}{2} & 2 \end{bmatrix} \cdot \begin{bmatrix} \frac{1}{3} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0.156 \\ 0.0938 \\ 0.125 \end{bmatrix}$$

This means that, when calculating the steady-state norm, the weight individual 7 assigns to individual 1 is equal to 0.156. Let us see how it is calculated. We have:

$$q_{71}^{\infty} = \underbrace{0.25}_{e_1(C_1)}(\underbrace{\frac{15}{8}}_{m_{77}} \times \underbrace{\frac{1}{3}}_{g_{71}} + \underbrace{\frac{9}{8}}_{m_{78}} \times \underbrace{0}_{g_{81}} + \underbrace{1}_{m_{79}} \times \underbrace{0}_{g_{91}}) = 0.156$$

Thus, the positive weight of 0.156 that individual 7 assigns to 1 is due to the fact that there is a direct link between 7 and 1. Observe that the other individuals 8 and 9 also assign a positive weight to individual 1, albeit smaller, because of the direct link between between 7 and 1. We can then perform the same exercise to determine how the weights are determined between individuals $7, 8, 9$ and individuals in C_2. Clearly, it will be through the link between 8 and 6. So, it is clear here how the position in the network of each individual has an impact on the weight that will be assigned in the steady-state social norm.

To see that, let us now determine how social norms converge in steady state. Assume the following (transpose) vector of initial norms:

$$[\mathbf{s}^{(\mathbf{0})}]^T = \begin{bmatrix} 0.1 & 0.2 & 0.3 & 0.4 & 0.5 & 0.6 & 0.7 & 0.8 & 0.9 \end{bmatrix} \tag{B.10}$$

It is then easily verified that the steady-state (transpose) vector of norms is given by:

$$[\mathbf{s}^{(\infty)}]^T = \begin{bmatrix} 0.225 & 0.225 & 0.225 & 0.5 & 0.5 & 0.5 & 0.328 & 0.397 & 0.363 \end{bmatrix} \tag{B.11}$$

Indeed, all agents belonging to the first closed communication class $C_1 = \{1, 2, 3\}$ converge to the same norm equal to 0.225, all agents belonging to the second closed communication class $C_2 = \{4, 5, 6\}$ converge to the same norm equal to 0.5, and, finally, all the other agents $7, 8, 9$ converge to different steady-state norms given by 0.398, 0.397 and 0.363, respectively. Even though individuals $7, 8, 9$ have high initial assimilation norms, they end up converging to a much lower level of assimilation norm than individuals belonging to C_2 because their weights are zero and their steady state norm is a convex combination of the norms of the two closed communication classes, which have norms between 0.225 and 0.5.

Let us now study the dynamics of the assimilation norms of individuals 2, 4, and 9, who belong to different groups. It is depicted in Figure B.3.

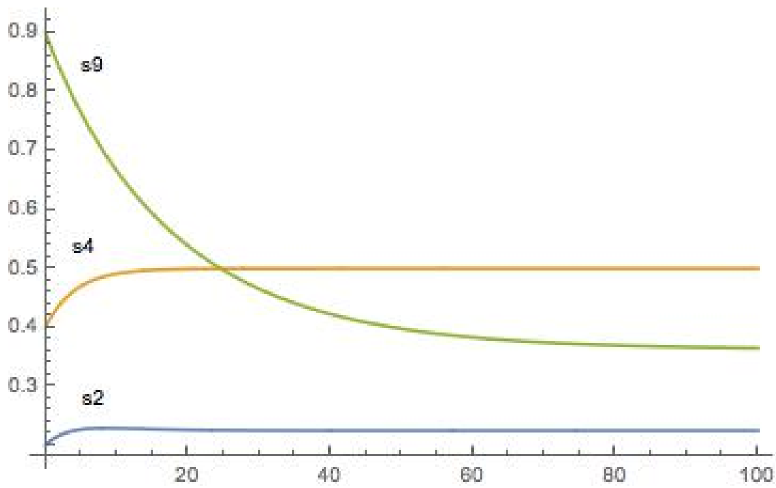

Figure B.3: Norm convergence with two communication classes

The assimilation norm of individual 2, who belongs to the first communication class C_1, converges to a low-value of assimilation (0.225) because all agents in this communication class have a low initial norm in terms of assimilation. The same intuition applies for the second closed communication class C_2 but with a higher value of the assimilation norm because of higher initial norms. Finally, for individuals 7, 8 and 9, who do not belong to any closed communication class, their steady-state norm is totally independent of their initial norms, which have high values. Their steady-state norms only depend on their position in the network, the distance between their location, and that of the two closed communication classes. We can see from network 3 in Figure 3 that there is always a path between any agent 7, 8, 9 and the two communication classes. As a result, the limit norm will be a combination of the steady-state norms of the two closed communication classes.

C Proof of equation (14)

Recall that

$$\mathbf{x}^t = \frac{1}{1+w}[\mathbf{I} - \theta\mathbf{G}]^{-1}\mathbf{s}^t$$

Then

$$\mathbf{x}^t - \mathbf{s}^t = \Big\{\frac{1}{1+w}[\mathbf{I} - \theta\mathbf{G}]^{-1} - \mathbf{I}\Big\}\mathbf{s}^t$$

Notice that the matrix $\mathbf{M}$ is a stochastic matrix. Denote $\mathbf{B} = \big\{\frac{1}{1+w}[\mathbf{I} - \theta\mathbf{G}]^{-1} - \mathbf{I}\big\}$. Then, $b_{ii} = m_{ii} - 1 < 0$, $b_{ij} = m_{ij} > 0$, for all $i \neq i$, and $\sum_{j\neq i} b_{ij} < 1$. Moreover, $b_{ii} = -\sum_{j\neq i} b_{ij}$

Looking at the ith row of $\mathbf{x}^t - \mathbf{s}^t$, we get

$$x_i^t - s_i^t = \sum_{j \neq i} m_{ij} s_j^t - \sum_{j \neq i} m_{ij} s_i^t = \sum_{j \neq i} m_{ij} (s_j^t - s_i^t)$$

D Speed of convergence

It is well known that, for a Markov process with transition matrix $\mathbf{T}$, the speed of convergence is inversely related to the second largest eigenvalue of $\mathbf{T}$, i.e. $\lambda_2(\mathbf{T})$. Moreover, as shown by Golub and Jackson (2012), if agents are divided into groups, each starting with a given norm, $\lambda_2(\mathbf{G})$ is a measure of homophily, which they called *spectral homophily*.[5] In our case, however, the speed of convergence is determined by $\lambda_2(\mathbf{T})$, which is different from $\lambda_2(\mathbf{G})$. Therefore, we would like to study how these two second largest eigenvalues relate to each other.

PROPOSITION **10** *Consider matrices* $\mathbf{G}$ *and* $\mathbf{T}$. *Then,*

$$\lambda_2(\mathbf{T}) = \frac{\gamma}{1 + \omega[1 - \lambda_2(\mathbf{G})]} + 1 - \gamma, \tag{D.1}$$

Moreover, the dynamics of norms and actions using the $\mathbf{T}$ *matrix is faster than that of the one using* $\mathbf{G}$ *if and only if* $\omega > \frac{1}{\lambda_2(\mathbf{G}) - (1-\gamma)}$.

Depending on the taste for conformity ω and on the parameter γ, this proposition shows that the convergence of social norms can be faster in our model than under the standard updating rule (Golub and Jackson (2010)) using the matrix $\mathbf{G}$.[6] To understand this, consider two extreme cases: $\gamma = 0$ and $\gamma = 1$.

When $\gamma = 0$, then $\mathbf{s}^{t+1} = \mathbf{s}^t$ (see equation (4)), agents perfectly adapt their own norms to their past norms, and no dynamics will take place. As a result, by continuity, we can say that for extremely low levels of γ, our dynamics is much slower than that of the standard updating with $\mathbf{G}$.

On the contrary, when $\gamma = 1$, $\mathbf{s}^{t+1} = \mathbf{x}^t$ and agents do not take into account their past norms but just try to be consistent with their past actions. This opens up the possibility

[5] In fact, the quantity $1 - |\lambda_2(\mathbf{G})|$ is called the *absolute spectral gap* and is an important measure of the system tendency to equilibrate.

[6] Observe that different closed communication classes have different speeds of convergence, which means that different groups of agents can converge to their steady-state norms at different speeds. To calculate the speed of convergence of each closed communication class C_p, one needs to compute the second largest eigenvalue of the submatrix of the communication class, $\lambda_2(\mathbf{C}_p)$.

of a faster process. In particular, our process is faster than under the standard updating rule if and only if $\omega > 1/\lambda_2(\mathbf{G})$. Indeed, with consistency, agents' norms get closer to their actions. Since the equilibrium actions of agents are proportional to their centrality, then, if the spectral homophily is low (i.e. low $\lambda_2(\mathbf{G})$), the simple updating on $\mathbf{G}$ is relatively fast. Then, our dynamics is faster if agents choose to conform to their neighbors' choices. This happens if the taste for conformity ω is high.

Finally, when $\gamma \in (0, 1)$, we have mixed effects. However, the proposition shows that, in order to have a fast convergence, the taste for conformity ω has to be large enough compared to spectral homophily.

To gain some additional intuition, consider the three networks in Figure 3, and the dynamics of the social norms in each network displayed in Figures B.1, B.2, and B.3 (in Appendix B), respectively. These figures show that network 3 experienced the fastest convergence (see Figure B.3) while, individuals in network 2 (Figure B.2) experience the slowest convergence of their social norms. Since, in these three figures, the parameters ω and γ are hold constant, the differences in the speed of convergence are only due to the differences in $\lambda_2(\mathbf{G})$. We have:

$$\lambda_2(\mathbf{G_3}) = 0.7675 < \lambda_2(\mathbf{G_1}) = 0.8108 < \lambda_2(\mathbf{G_2}) = 0.8791$$

Indeed, the order of the inverse of the second largest eigenvalues reflects the order in the speed of convergence.

Proposition 10 shows that, apart from the second largest eigenvalue, a change in the conformism parameter ω and the updating parameter γ also affect the speed of convergence for a given network structure. Let us focus on network 1. In the dynamics displayed in Figure B.1, we use $\gamma = 0.5$ and $\omega = 0.5$. In the dynamics displayed in Figure D.1 below, we take instead $\gamma = 0.5$ and $\omega = 0.3$ (left panel) and $\gamma = 0.6$ and $\omega = 0.5$ (right panel). In other words, in the former, we only decrease ω while, in the latter, we only increase γ.

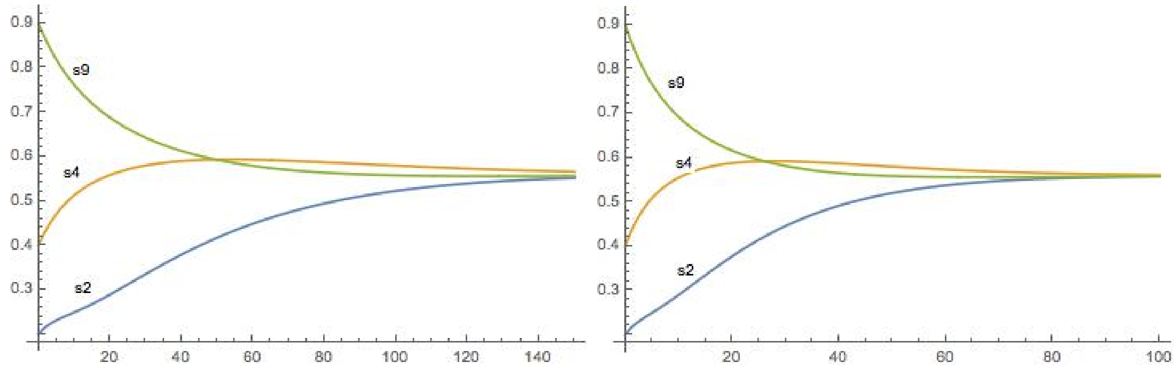

Figure D.1: Convergence of norms in network 1 when $\gamma = 0.5$ and $\omega = 0.3$ (left panel) and $\gamma = 0.6$ and $\omega = 0.5$ (right panel)

Consider first a change in ω. By equation (**??**), an increase of w would decrease $\lambda_2(\mathbf{T})$ thus making the process faster. In our simulations, we decrease ω so that the dynamics is slower. Indeed, a lower taste for conformity ω means a lower tendency for agents to conform to others' social norms. This means that individuals have less incentives to follow others' behaviors and, through updating, having norms relating to each other. This implies that the process is going to be slower. Consider now an increase of γ. Then, during the updating process, the agents weigh more past actions (consistency) but since actions relate to others' actions and norms, this larger weight makes the norms more interconnected and reach a steady state faster. This is why the convergence is quicker in the right panel of Figure D.1.

E Peer definitions

In this section, we consider two alternative definitions of peers.

E.1 Peers as indegrees

We now define a peer as someone who nominates you but not necessary the other way around. In other words, being a peer is now a measure of *popularity* rather than a measure of *role model*. So, if individual i nominates individual j but the reverse is not true, then $g_{ji} = 1$ and $g_{ij} = 0$. If we consider network 3 in Figure 3, for *indegrees*, the adjacency matrix is not anymore given by (B.3) but by:

$$\mathbf{G_3} = \begin{bmatrix} 0 & 0 & \frac{1}{2} & 0 & 0 & 0 & \frac{1}{2} & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{2} & 0 & \frac{1}{2} & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{3} & \frac{1}{3} & 0 & 0 & \frac{1}{3} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} & 0 \end{bmatrix} \tag{E.1}$$

It is easily verified that, in network 3, there is now only one closed communication class $C_1 = \{7, 8, 9\}$ since $g_{17} > 0$ and $g_{68} > 0$ while $g_{17} = 0$ and $g_{68} = 0$. This is exactly the opposite of the case when peers were defined as outdegrees where there were two closed communication classes, $C_1 = \{1, 2, 3\}$ and $C_2 = \{4, 5, 6\}$, and individuals $7, 8, 9$ did not belong to any closed communication class. Now, when peers are indegrees, individuals $1, 2, 3, 4, 5, 6$ do not belong to any closed communication class, so that $m = 1$, $n_1 = 3$ and $n_Q = 6$. The canonical form of $\mathbf{G_3}$ is given by $\widetilde{\mathbf{G}}_3$, where

$$\mathbf{C_1} = \begin{bmatrix} 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & 0 \end{bmatrix}; \mathbf{Q} = \begin{bmatrix} 0 & 0 & \frac{1}{2} & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & 0 & 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} & 0 \end{bmatrix}; \mathbf{Q_{C_1}} = \begin{bmatrix} \frac{1}{2} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

Let us now determine the steady state norms of all individuals in the network. Since there is one closed communication class and six individuals who do not belong to any closed communication class, there will be one social norm in steady state. We obtain:

$$\mathbf{T}_{\mathbf{3}}^{\infty} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0.333 & 0.333 & 0.333 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0.333 & 0.333 & 0.333 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0.333 & 0.333 & 0.333 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0.333 & 0.333 & 0.333 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0.333 & 0.333 & 0.333 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0.333 & 0.333 & 0.333 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0.333 & 0.333 & 0.333 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0.333 & 0.333 & 0.333 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0.333 & 0.333 & 0.333 \end{bmatrix}$$

The results in terms of assimilation are exactly the opposite to the ones obtained when peers were defined as outdegrees (see (B.9)) since, in the latter, individuals $1, 2, 3, 4, 5, 6$ (resp. individuals $7, 8, 9$) were the most (less) likely to be assimilated while, in the former, they are the less (most) likely to be assimilated. Indeed, if we again assume the same vector of initial norms given by (B.10), then, when peers are defined as indegrees, all individuals will converge to a social norm equal to: $s^{\infty} = 0.8$. This norm is quite high because all individuals in the closed communication class have a very high initial social norm (between 0.7 and 0.9). When peers were defined as outdegrees, we showed that the long-run assimilation norms were given by (B.11). In particular, individuals 7, 8 and 9 had low long-run social norms (less than 0.4) because their initial norms did not matter at all and only their position in the network, which determined their connection to the two closed communication classes, had an impact on the determination of their long-run norms.

E.2 Peers as mutual friends

We now define peers as *mutual friends* in two different ways. First, we say that there is a link between individuals i and j i.e. $g_{ij} = 1$ so that i and j are peers, if *either* i has nominated j *or* j has nominated i. This is the standard way researchers have been defining peers in empirical works. Second, we define a link between two individuals i and j if *both* individual i nominates individual j and j nominates i. This is a more restrictive way of defining peers.

Let us start with the first definition and consider network 3 in Figure 3. We have:

$$\mathbf{G_3} = \begin{bmatrix} 0 & 0 & \frac{1}{2} & 0 & 0 & 0 & \frac{1}{2} & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{2} & 0 & \frac{1}{2} & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{3} & \frac{1}{3} & 0 & 0 & \frac{1}{3} & 0 \\ \frac{1}{3} & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{3} & \frac{1}{3} \\ 0 & 0 & 0 & 0 & 0 & \frac{1}{3} & \frac{1}{3} & 0 & \frac{1}{3} \\ 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} & 0 \end{bmatrix}$$

It should be clear that the network is now *strongly connected* and thus all individuals will converge to the same social norm and will all exert the same assimilation effort.

Consider now the second definition of peers where both need to nominate each other to be considered as peers. We have:

$$\mathbf{G_3} = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{2} & 0 & \frac{1}{2} & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} & 0 \end{bmatrix}$$

In this case, there are three closed communication classes, $C_1 = \{1, 2, 3\}$, $C_2 = \{4, 5, 6\}$ and $C_3 = \{7, 8, 9\}$ so that $m = 3$, $n_1 = n_2 = 3$ and $n_Q = 0$. There will be three different social norms, which will mostly depend on the initial social norms so that the social norm of C_1

will be lower than that of C_2, which, in turn, will be lower than C_3. We indeed obtain:

$$\mathbf{T}_{\mathbf{3}}^{\infty} = \begin{bmatrix} 0.25 & 0.25 & 0.5 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0.25 & 0.25 & 0.5 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0.25 & 0.25 & 0.5 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.333 & 0.333 & 0.333 & & 0 & 0 \\ 0 & 0 & 0 & 0.333 & 0.333 & 0.333 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.333 & 0.333 & 0.333 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0.333 & 0.333 & 0.333 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0.333 & 0.333 & 0.333 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0.333 & 0.333 & 0.333 \end{bmatrix}$$

Using the same initial norms defined by (B.10), we easily get:

$$[\mathbf{s}^{(\infty)}]^T = \begin{bmatrix} 0.225 & 0.225 & 0.225 & 0.5 & 0.5 & 0.5 & 0.8 & 0.8 & 0.8 \end{bmatrix}$$

E.3 Some more general results

Let us try to obtain some general results about the definition of peers. Consider the first network in Figure E.1 and define peers as outdegrees. Then, it is clear that agents 4 and 5 are the *influencers* of the whole dynamics, since they belong to the only closed communication class, while agents 1, 2 and 3 are the *influenced* individuals since they do not belong to any closed communication class. Then, if we change the definition of links from outdegrees to indegrees, we have the opposite result, namely the influencers become the influenced. If we now consider reciprocal nomination (peers as mutual friends), then the two groups consisting of agents 1, 2, and 3, and of agents 4 and 5, are disconnected classes without interaction, and no convergence happens.

More generally, consider the definition of peers as indegrees and outdegrees. Consider the original adjacency matrices $\bar{\mathbf{G}}_{out}$ and $\bar{\mathbf{G}}_{in}$ (i.e the adjacency matrices that are *not* row-normalized) and the *diagonal* matrices $\mathbf{D}_{out}$ and $\mathbf{D}_{in}$ of indegrees and outdegrees, respectively, that is the diagonal of $\mathbf{D}_{out}$ ($\mathbf{D}_{in}$) contains the number of outdegree (indegree) links of each node while its off-diagonals have zeros. Then, by definition

$$\mathbf{G}_{out} := \mathbf{D}_{out}^{-1}\bar{\mathbf{G}}_{out} \quad \text{and} \quad \mathbf{G}_{in} := \mathbf{D}_{in}^{-1}\bar{\mathbf{G}}_{in}$$

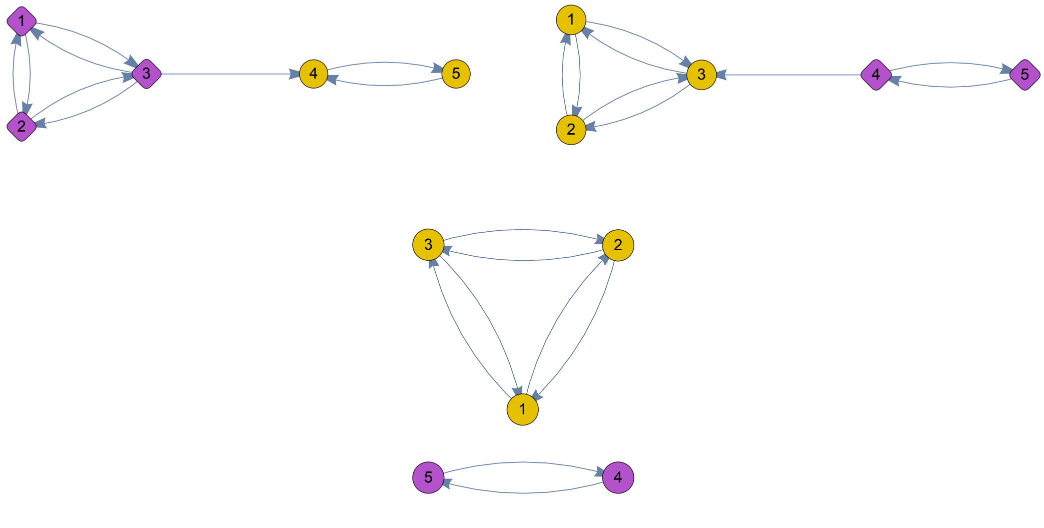

Figure E.1: Peers definition: Outdegree, Indegree, Reciprocal

We then have the following relationship:

$$\mathbf{G}_{in} = \mathbf{D}_{in}^{-1}\left(\mathbf{D}_{out}\mathbf{G}_{out}\right)^{T} \tag{E.2}$$

However, without further restrictions, there is no way to relate steady-state norms and actions in the two cases, since there is not a clear relationships between the eigenvectors of a matrix and of the same matrix times a non-scalar diagonal matrix. As a general result we can only state the following:

Remark **1** *When peers are defined either as indegrees or outdegrees, the convergence levels and properties are different but the speed of convergence of social norms and actions is the same.*

This result immediately follows from the fact that a matrix and its transpose share the same eigenvalues. Intuitively, given that the second largest eigenvalue measures how the network can be splitted into two parts, changing the direction of all the links does not affect this property and, thus, the speed of convergence is the same. However, it clearly affects the convergence properties.

Consider again the networks in Figure E.1 and assume that the initial norms are given by: $s_1^0 = 0.1$, $s_2^0 = 0.2$, $s_3^0 = 0.3$, $s_4^0 = 0.4$, and $s_5^0 = 0.5$. It is easily seen that, in the first network, agents 4 and 5 form a unique closed communication class, so that, for all i,

$s_i^\infty \in [0.4, 0.5]$. In the second network, agents 1,2, 3 form a unique closed communication class so that, for all i, $s_i^\infty \in [0.1, 0.3]$. As a result, in both cases, a consensus (convergence of social norms) will emerge but to very different levels. In the third network, there are clearly two closed communication classes and thus the first three agents converge to the same social norm of assimilation in $[0.1, 0.3]$ while the last two agents will have norms converging to a long-run norm belonging to $[0.4, 0.5]$.

F Proofs of the results in the Online Appendix

Proof of Proposition 7: Consider the case in which $\mathbf{G}$ is aperiodic. Then, as shown in Proposition 1, $\lim_{t\to+\infty} \mathbf{T}^t = \lim_{t\to+\infty} \mathbf{G}^t$. A strongly connected network implies $\mathbf{G}$ to be irreducible. As shown, for example, in Golub and Jackson (2010), for an irreducible and aperiodic network,

$$\lim_{t\to+\infty} \mathbf{G}^t = \begin{pmatrix} \mathbf{e}(\mathbf{G}) \\ \cdots \\ \cdots \\ \mathbf{e}(\mathbf{G}) \end{pmatrix}$$

Consider now the case in which $\mathbf{G}$ is periodic. Proposition 1 states that $\lim_{t\to+\infty} \mathbf{T}^t = \lim_{t\to+\infty} \mathbf{G}_\epsilon{}^t$. Then $\mathbf{G}_\epsilon$ is aperiodic. For the same reasoning as before we get

$$\lim_{t\to+\infty} \mathbf{G}_\epsilon{}^t = \begin{pmatrix} \mathbf{e}(\mathbf{G}_\epsilon) \\ \cdots \\ \cdots \\ \mathbf{e}(\mathbf{G}_\epsilon) \end{pmatrix}$$

This proves the result. ∎

Proof of Proposition 8 Consider first the agents belonging to the closed communication class. The dynamics of their norms is represented by the matrix $\mathbf{C}$. Previous proposition characterizes the convergence for the norms of these agents. Consider now also agents not belonging to the closed communication class. The entire process is a uni-reducible Markov process and the steady state for all agents is the steady state of agents belonging to the closed communication class. ∎

Proof of Proposition 9: Consider first agents belonging to closed communication class.

Then, by previous analysis it is proved that (B.6) holds.
Consider now the weight assigned by each agent in Q to any other in Q. It is a well-known results that the matrix of these weights is null.[7] This immediately comes from the fact that, taking the power matrix $\mathbf{G}^t$, the block corresponding to $\mathbf{Q}$ is given by $\mathbf{Q}^t$. Since $\mathbf{Q}$ is a substochastic matrix, then $\mathbf{Q}^\infty = \mathbf{O}$.
Consider now how each $\mathbf{g}_{i_p}^\infty$ is shaped. Recall, first, that all agents belonging to a communication class C_p have norms converging to the same value. Call x_{kp} the weight that agent $k \in Q$ assigns in steady state to the communication class C_p. Then $\mathbf{x_p}$ is the vector of dimension n_q assigning the steady state weights of all agents in Q to the communication class C_p. As proved in Theorem 4.4 of Senata (1981), we have:

$$\mathbf{x_p} = [\mathbf{I} - \mathbf{Q}]^{-1}\mathbf{Q}_{\mathbf{C_p}}\mathbf{1} \tag{F.1}$$

Then, if we want to determine the weight that agent $i \in C_p$ contributes to $\mathbf{x_p}$, it is enough to multiply $\mathbf{x_p}$ by $e_i(\mathbf{C_p})$. Indeed, we have:

$$\mathbf{g}_{i_p}^\infty = e_i(\mathbf{C_p})[\mathbf{I} - \mathbf{Q}]^{-1}\mathbf{Q}_{C_p}\mathbf{1},\ \forall i \in C_p,\ \forall p = 1, \ldots, m \tag{F.2}$$

This completes the proof. ∎

Proof of Proposition 10 Consider a generic matrix $\mathbf{A}$ and its vector of eigvenvalues $\boldsymbol{\lambda}(\mathbf{A})$, with $\lambda_1(\mathbf{A}) > \lambda_2(\mathbf{A}) > ....$ Consider a polynomial function $P(\cdot)$. Then, it is well-known that $\boldsymbol{\lambda}(P(\mathbf{A})) = P(\boldsymbol{\lambda}(\mathbf{A}))$. Consider now $\mathbf{T} \equiv \frac{\gamma}{(1+w)}\mathbf{M}(\theta, \mathbf{G}) + (1-\gamma)\mathbf{I}$, and recall that

$$\mathbf{M}(\theta, \mathbf{G}) = \sum_{k=0}^{\infty} \theta^k \mathbf{G}^k. \tag{F.3}$$

Then,

$$\boldsymbol{\lambda}(\mathbf{M}(\theta, \mathbf{G})) = \sum_{k=0}^{\infty} \theta^k \boldsymbol{\lambda}^k(\mathbf{G}). \tag{F.4}$$

Notice that $\sum_{k=0}^{\infty} \theta^k \boldsymbol{\lambda}^k(\mathbf{G})$ preserves the order of the eigenvalues. Then

$$\lambda_2(\mathbf{M}(\theta, \mathbf{G})) = \sum_{k=0}^{\infty} \theta^k \lambda_2^k(\mathbf{G}). \tag{F.5}$$

Observe that, $\theta\lambda_2(\mathbf{G}) < \theta\lambda_1(\mathbf{G}) < 1$, where the last inequality holds by assumption. Then

[7] See, for example, Theorem 4.3 in Senata (1981).

$$\lambda_2(\mathbf{M}(\theta, \mathbf{G})) = \frac{1}{1 - \theta\lambda_2(\mathbf{G})}, \tag{F.6}$$

and,

$$\lambda_2(\frac{\gamma}{1+\omega}\mathbf{M}(\theta, \mathbf{G})) = \frac{\gamma}{(1+\omega)(1 - \theta\lambda_2(\mathbf{G}))}. \tag{F.7}$$

Since $\theta = \frac{\omega}{1+\omega}$, we have:

$$\lambda_2(\frac{\gamma}{1+\omega}\mathbf{M}(\theta, \mathbf{G})) = \frac{\gamma}{[1 + \omega(1 - \lambda_2(\mathbf{G})]}. \tag{F.8}$$

Recall that we have shown that: $\mathbf{MT} = \mathbf{TM}$. This implies that $\frac{\gamma}{1+\omega}\mathbf{M}$ and $\mathbf{T}$ commute and, thus, share the same eigenvectors. Denote by $\mathbf{v}_2$ the eigenvector associated to the second largest eigenvalue of $\frac{\gamma}{1+\omega}\mathbf{M}$. Then, by definition,

$$\mathbf{M}\mathbf{v}_2 = \lambda_2(\mathbf{M})\mathbf{v}_2. \tag{F.9}$$

With some algebra, we can show that:

$$\frac{\gamma}{1+\omega}\mathbf{M}\mathbf{v}_2 = \frac{\gamma}{1+\omega}\lambda_2(\mathbf{M})\mathbf{v}_2, \tag{F.10}$$

and

$$\big[\frac{\gamma}{1+\omega}\mathbf{M} + (1-\gamma)\mathbf{I}\big]\mathbf{v}_2 = \big[\frac{\gamma}{1+\omega}\lambda_2(\mathbf{M}) + (1-\gamma)\big]\mathbf{v}_2. \tag{F.11}$$

Since $\mathbf{T} = \frac{\gamma}{1+\omega}\mathbf{M} + (1-\gamma)\mathbf{I}$, we have:

$$\lambda_2(\mathbf{T}) = \frac{\gamma}{[1 + \omega(1 - \lambda_2(\mathbf{G})]} + (1-\gamma) \tag{F.12}$$

Since the speed of convergence is inversely related to the second largest eigenvalue, the dynamics ruled by $\mathbf{T}$ is faster than the one ruled by $\mathbf{G}$ if and only if $\lambda_2(\mathbf{T}) < \lambda_2(\mathbf{G})$. With simple algebra, we obtain the result stated in the proposition. ▮

References

Seneta, E. (1981), *Non-Negative Matrices and Markov Chains*, Second Edition, New York: Springer.