Why Do Homogeneous Friendships Persist in a Diverse Population?

Making Sense of Homophily

Firman M. Firmansyah¹ and Ahmad R. Pratama²

¹Department of Technology and Society, Stony Brook University

²Department of Informatics, Universitas Islam Indonesia

Author Note

Corresponding concerning this article should be addressed to Firman M. Firmansyah, Department of Technology of Society, 1432 Computer Science, Stony Brook University, Stony Brook, NY 11794-4404. Email: manda.firmansyah@stonybrook.edu. Twitter: @firmansm.

Abstract

Homophily is one of the robust findings in social network studies. It persists even in a diverse population where the opportunity to develop homogeneous friendship is not greater than a mere chance and the process of developing heterogeneous friendship is facilitated. In this study, we introduce the Framework for Intergroup Relations and Multiple Affiliations Networks (FIRMAN), derived from social identity theory and social network framework, that can explain why that is the case. We begin by explaining its components: social identity space, social identity distance, length of ties, tie outreachability, and tie capacity. Then, through 7000 agent-based simulations, we demonstrate how the interaction of tie outreachability and tie capacity constraints heterogeneous friendship formations, which in turn make homophily inevitable even in a very diverse population. Surprisingly, the presence of even a small percentage ($< 15\%$) of agents who can only develop homogeneous friendships can affect the whole population, preventing other agents from developing heterogeneous friendships. We conclude by providing some directions for future research.

Keywords: homophily, friendship, social identity, social network, diversity

Why Do Homogeneous Friendships Persist in a Diverse Population?

Making Sense of Homophily

Homophily, the tendency for friendships to develop among homogeneous individuals (Lazarsfeld & Merton, 1954), is one of the robust findings in social network studies (McPherson et al., 2001). It applies not only for salient aspects such as gender (Laniado et al., 2016), race (Campigotto et al., 2021), ethnicity (Wölfer & Hewstone, 2018), and weight (Simpkins et al., 2013), but also for less salient ones such as values (Dehghani et al., 2016), personality (Ilmarinen et al., 2017), academic performance (Smirnov & Thurner, 2017), and genes (Christakis & Fowler, 2014). Multiple studies that were conducted in various contexts (Hofstra et al., 2017; Mazur & Richards, 2011) and across cultures (French et al., 2012; Leszczensky & Pink, 2017; Smith et al., 2016) have yielded the same conclusion: homophily is a universal phenomenon.

Past research highlights structural factors such as propinquity (Kruse et al., 2016; Mouw & Entwisle, 2006), focus closure (Kossinets & Watts, 2009), and triadic closures (Asikainen et al., 2020) that give homophily more rooms to grow. In this respect, homophily is the direct consequence of the fact that people of the same group tend to live in the same region, go to a nearby place, or do similar activities. Thus, individuals have little to no choice but to develop friendships with those who are alike. Other research underlines psychological aspects such as personal preferences (Anderson et al., 2014; Yu & Xie, 2017), similarity attractions (Ilmarinen et al., 2016; Stark & Flache, 2012), and intergroup attitudes (Bahns et al., 2015; Fischer, 2011; Poteat et al., 2011) that make homogeneous friendships more desirable. In typical instances, individuals intentionally choose to befriend similar others despite having an opportunity to flock with different others.

Based on the antecedents, homophily is thus distinguished as either induced homophily (driven by structural factors) or choice homophily (driven by personal preference) (J. M. McPherson & Smith-Lovin, 1987). Yet, it is still not easy to identify which homophily falls into which category. In induced homophily, for example, some individuals do prefer to flock with similar others, which is also why they go to a place or live in a region where they will be surrounded by others alike in the first place. Some scholars tried to disentangle these antecedents using various techniques including conditional logit model (Zeng & Xie, 2008), exponential random graph

model (Wimmer & Lewis, 2010), revealed preference theory (Currarini et al., 2010), and simulations (Avin et al., 2020). However, they hardly arrived at the same conclusion fueling a prolonged debate on which antecedent (structural vs preference) is more dominant in causing homophily (Lawrence & Shah, 2020).

Regardless of the debate, it makes sense to assert that once the structural constraints have been eliminated and the psychological aspects have been intervened, homogeneous friendships should not develop more than a mere chance. In other words, heterogeneous friendship should have, at least, an equal opportunity to blossom when diverse individuals meet and do things together in the same intimate place. Nevertheless, this notion is too good to be true. In a study of racially diverse newcomers in a university, most students still developed homogeneous friendships even though they were nudged to do the opposite (Mollica et al., 2003). In racially mixed schools, most students still tended to have lunch together with those of the same race (Ramiah et al., 2015). Even when such further interventions have been conducted (i.e., assigning diverse students to engage in small group activities), heterogeneous friendships did not last long (Boda et al., 2020).

In this paper, we introduce a theoretical framework called the Framework for Intergroup Relations and Multiple Affiliations Networks (FIRMAN) that can explain why individuals still develop friendships with others alike despite having an opportunity to flock with different others. We begin by first, introducing its components, namely social identity space (abstract space constituted by multiple social identities), social identity distance (geometric distance between individuals in social identity space), length of ties (geometric length of ties connecting nodes in social identity space), tie outreachability (maximum geometric *length* of ties that a node can generate in social identity space), and tie capacity (maximum *number* of ties that a node can generate and maintain at a time). Then, through a series of agent-based simulations, we demonstrate how this framework elucidates the emergence of homophily in diverse populations.

Framework for Intergroup Relations and Multiple Affiliations Networks (FIRMAN)

Suppose there are n individuals i in a hypothetical population. They have the same j number of social identities SI_j derived from the groups to which they belong (Hogg, 2016; Tajfel & Turner, 1979). Let $n = 8$, $i \in \{1 \dots 8 \mid 1 = A, \dots, 8 = H\}$, and $j = 2$, $j \in \{1, 2\}$. Let *shape* depict the first social identity (SI_1) and *color* depict the second social identity (SI_2). For the sake of parsimony,

each social identity has two groups only: round and square for SI_1 (shape), white and black for SI_2 (color). Let SI_{i1} = square for $i \in \{1 \dots 4\}$ and SI_{i1} = round for the rest. Let SI_{i2} = white for $i \in \{1, 2, 5, 6\}$ and SI_{i2} = black for the rest. White and round are arbitrarily assigned to zero (0); black and square to one (1). Thus, each social identity follows the Bernoulli distribution with a binary value of k and a probability of p denoting the proportion of the minority within the population, leading to another probability of $p-1$ for the majority group as shown in equation (1). This distribution reaches its maximum variance at $p = .5$, in which the notion of majority and minority becomes irrelevant as there are two distinct groups with equal size, instead.

$$\forall SI; f(k; p) = p^k(1-p)^{1-k} \text{ for } k \in \{0, 1\} \quad (1)$$

Components

Social Identity Space

Suppose sets of individuals I_i and social identities SI_j constitute a social identity matrix $SI_{i \times j}$. Using this matrix, we can place i individuals into *social identity space* as shown in Figure 1. This space has j dimensions corresponding to the number of columns of the $SI_{i \times j}$ matrix, which in this case is two. The axes are intentionally drawn with a different scale (weight) to visualize the assumption that individuals of different colors (SI_2) are more distant than those of different shapes (SI_1). A real-life example: people of different political ideologies have been more distant than those of different races (Iyengar & Westwood, 2015).

Figure 1

A Two-Dimensional Social Identity Space

Note. A two-dimensional social identity space derived from $SI_{8 \times 2}$ matrix. Multiple similarities put individuals into the same coordinates, while differences put them in distant coordinates.

We call this space *social identity space* because it is constituted by social identities. The name also distinguishes itself from a rather related framework, Blau space (J. M. McPherson & Ranger-Moore, 2017), which uses socio-demographic variables as the dimensions (M. McPherson, 2004). As past studies revealed, social identities can be derived from arguably anything that is shared, including trivial attributes such as name-tag color (Jackson et al., 2018). In other words, it does not have to be derived from meaningful socio-demographic variables. If the shared attribute binds individuals together (Turner et al., 1994), creates boundaries between ingroup and outgroup members (Hogg & Turner, 1987), and directs them to relatively similar collective attitudes and behaviors (van Zomeren et al., 2008), it can be regarded as a social identity, and thus be added as a dimension in the social identity space.

Social Identity Distance

By having individuals placed in the social identity space, we can visually calculate the *social identity distance* between them. It reflects the degrees of similarities or differences between a pair of individuals. We use the Manhattan distance to calculate social identity distance to maintain discrete values. As in this case social identities are weighted, we also add weights ($w_u, u \in j, j \in \{1, 2\}$) to calculate the distance as follows.

$$w_dist_{xy} = W_u \cdot \sum |SI_{xj} - SI_{yj}|; x, y \in I \quad (2)$$

We assign values $w_1 = 1$ for SI_1 and $w_2 = 2$ for SI_2 . Even though these values are arbitrary, doing so gives a further benefit in interpreting the social identity distance between individuals. In Figure 1, zero social identity distance ($w_dist_{xy} = 0$) denotes that two individuals are of the same shape and color (e.g., A and B); one social identity distance ($w_dist_{xy} = 1$) denotes that they are of different shapes but the same color (e.g., A and E); two social identity distance ($w_dist_{xy} = 2$) denotes that they are of the same shape but the different colors (e.g., A and C); three social identity distance ($w_dist_{xy} = 3$) denotes that they are of different shapes and different colors (e.g., A and G).

Social identity distance in this framework is different from Bogardus' (1947) notion of social distance, which reflects the extent to which one is willing to engage in social contacts with members of different groups. As implied in the items of his scale (Wark & Galliher, 2007), social

distance ranges from the closest form of intergroup inclusions (e.g., kinship by marriage) to the farthest one (e.g., excluding different others from entering the country). Considering its subjective nature, individuals of the same group may perceive different social distances with the same outgroup members. For example, E, who is round-white, is hypothetically willing to accept members of the round-black group (i.e., G or H) as his close relatives' spouse, whereas F, who also happens to be round-white, can only accept members of the black-round group no more than just a visitor in the country. In this illustration, E has a closer Bogardus social distance with round-black members than F does.

In FIRMAN, on the other hand, the social identity distance between ingroup or outgroup members is fixed. Regardless of the differences between E and F in accepting round-black members (i.e., G or H) as closed contact, their social identity distance will remain the same, which is 2 (same shape, different color). It will change, however, if there are some changes in the social identity matrix. For instance, should E hypothetically turn to be black, its distance with G or H changes from 2 (same shape, different colors) to 0 (same shape, same color).

Length of Ties

Let nodes represent individuals and ties represent their relationships in the social identity space. As can be seen in Figure 2, the lengths of ties (LT) vary relative to the nodes' social identity distance. Ties between nodes with different social identities (i.e., $LT_{BC} = 2$; $LT_{CH} = 1$) are longer than those with the same social identities (i.e., $LT_{AB} = 0$; $LT_{EF} = 0$; $LT_{CD} = 0$).

Figure 2

A Hypothetical Network with Undirected Ties

Note. In this network, ties connecting nodes of the same social identities are shorter than ties connecting nodes of different social identities.

Variations in the length of ties, however, do not represent intimacy levels. They simply inform whether relationships between pairs of nodes in the space exist or not. Thus, even though the tie connecting EF ($LT_{EF} = 0$) is shorter than the tie connecting BC ($LT_{BC} = 2$), they have the same degree of intimacy of relationship in this framework. If intimacy levels are to be included (e.g., friendship vs acquaintanceship), it could be depicted by the thickness of ties following the notion of weak and strong ties (Granovetter, 1973). For now, for the sake of simplicity, let it be excluded from the equation.

Tie Outreachability

We argue that not all nodes can generate the same geometric length of ties in social identity space, which in turn, affects their ability to reach out to different others. Let us call this ability, *tie outreachability* (TO). In real life, tie outreachability analogues to the extent to which one can tolerate, accept, and navigate the conflicts perpetuated by social identity differences. As past studies show, heterogeneous relationships are more challenging and prone to conflicts compared to homogeneous relationships (Bratter & King, 2008; Heaton, 2002). In this framework, variations in tie outreachability represent the variations in individual tolerance and acceptance for differences (Hjerm et al., 2020). For example, some support interracial relationship but cannot get along with people of different political ideology (Iyengar & Westwood, 2015).

The range of tie outreachability in the population level corresponds to the range of the possible minimum and maximum values of social identity distance in the social identity space (in this example, it ranges from 0 to 3). This makes tie outreachability follows the binomial distribution with parameter trial m and probability q . The value of m is the same as the value of the maximum possible social identity distance in the space. The value of q , which denotes probability of success, will represent the variance of outreachability in the population level. Lower q represents a population overpopulated by nodes who can only generate shorter ties; higher q is otherwise. With infinite m and $q = .5$, the tie outreachability distribution will resemble the shape of a normal distribution.

$$TO, f(k; m, q) = q^k (1 - q)^{(1-k)} \text{ for } k = 2, m = \max(w_dist) \quad (3)$$

It is worth highlighting that tie outreachability is cumulative in nature. Individuals with higher tie outreachability can generate lengths of ties shorter than their maximum length of ties, but not vice versa. For example, an individual with $TO = 2$ can generate ties with $LT = 0$ but cannot generate ties with $LT = 3$. This logic is analogous to the fact that in real life, people who are engaged in intergroup relationships usually are also involved in intragroup relationships at the same time, but not vice versa.

Tie Capacity

We further argue that the numbers of ties that a node can generate and maintain at a time in social identity space are limited, just like the limited friendships that humans can make during their lifetimes (Dunbar, 2018). This ability, which we call *tie capacity*, will determine how dense a personal network can be. The higher the capacity, the more connections a node can potentially develop. In this framework, we use a normal distribution with parameter μ and standard deviation σ^2 , to generate tie capacity distributions at the population level as follows.

$$TC, f(x; \mu, \sigma^2) = 1 / \sigma (x - \mu / \sigma) \quad (4)$$

Friendship Development

Most individuals, if not all, need friends to fulfill their social needs (Hojjat & Moyer, 2017). Meanwhile, intimate meaningful friendships require reciprocity (Almaatouq et al., 2016; Harrigan & Yap, 2017). Under FIRMAN, friendships can be reciprocated and thus develop if and only if both individuals in the dyad can generate ties that can cover the social identity distance between them. This required condition will be satisfied if the minimum value of tie outreachability in a pair of nodes is greater than or equal to their weighted social identity distance as shown in equation (5).

$$friends_{xy} \leftrightarrow w_dist_{xy} \leq \min (TO_x, TO_y) \quad (5)$$

Equation (5) is especially relevant to the development of heterogeneous friendships that require longer ties. If one of the potential friends in a heterogeneous dyad can only generate ties shorter than the social identity distance, heterogeneous friendships cannot develop. For instance, interethnic friendships between immigrants and locals are rare, even in a melting pot, is not

because the former are lacking in social skills but because the latter too often did not reciprocate their friendship initiatives (Barwick, 2017). Meanwhile, homogeneous friendships blossom, last longer, and thus are more ubiquitous than heterogeneous ones (Kao & Joyner, 2004; Rude & Herda, 2010; Tulin et al., 2021) since, we argue, they require shorter ties that individuals with the same social identities can easily generate. Even more so, nodes of the same multiple social identities require zero length of ties.

While reciprocated ties are necessary, it is not sufficient in and of itself. Individuals cannot develop a new friendship, neither homogeneous nor heterogeneous, if their number of friends (num_f_i) has reached their maximum tie capacity. Number of friends is the sum of people either of the same or different group in the population, with whom an individual has friendship.

$$\forall x, y, friends_{xy} = \{(x,y)\}, num_f_x = \sum y, y \neq x; x, y \in I \quad (6)$$

Some individuals forgo a chance to develop intergroup friendship not because they hate different others, but simply because they already have enough friends (Pfister et al. 2020). Thus, both nodes in the dyad should at least have one tie available for friendship to develop. Combining these two axioms, therefore, friendship will develop if the following conditions are satisfied.

$$friends_{xy} \leftrightarrow w_dist_{xy} \leq \min (TO_x, TO_y) \wedge num_f_x < TC_x \wedge num_f_y < TC_y \quad (7)$$

Hypothetical Example

Let us use the hypothetical population presented in Table 1 to show how FIRMAN can illuminate homophily. This population is the same as the one introduced earlier with social identities converted into binary codes. It is worth highlighting that this population is very diverse as there are four different social identity intersection groups of the same size, so there is no majority nor minority among them. Let us assume that all individuals in the population have social needs, which can be fulfilled if they develop at least one friendship. All individuals can only make friends within this population. Individuals who receive friendship initiatives from others cannot and will not decline unless their conditions do not satisfy equation 7.

Table 1

Social Identities (SI), Tie Outreachability (TO), and Tie Capacity (TC) of Eight Individuals in a Hypothetical Population

Individual	SI ₁	SI ₂	TO	TC
A	0	1	3	3
B	0	1	2	2
C	1	1	2	3
D	1	1	1	1
E	0	0	1	1
F	0	0	0	2
G	1	0	0	3
H	1	0	1	2

The following is one of many possible friendship development scenarios. Suppose A ($TO_A = 3$, $TC_A = 3$) initiates a friendship attempt to G ($TO_G = 0$, $TC_G = 3$), $w_{\text{dist}_{AG}} = 3$, as shown in Figure 3(a). Since $LT_G = 0$, no matter how hard A tries to befriend G, its friendship initiative will not be reciprocated ($LT_G < w_{\text{dist}_{AG}}$). Instead, G develops a friendship with H ($TO_H = 1$, $TC_H = 2$), who is the only individual in the space that can be reached out given $LT_G = w_{\text{dist}_{AG}} = 0$. Realizing that its friendship initiative is not reciprocated, A tries to befriend E ($TO_E = 1$, $TC_E = 1$), $w_{\text{dist}_{AE}} = 1$, as shown in Figure 3(b). With $LT_E = w_{\text{dist}_{AE}} = 1$, their friendship could develop. However, it is too late since, by the time A approaches E, everyone in the space has simultaneously initiated a friendship with each other as shown in Figure 3(c). It includes F ($TO_F = 0$, $TC_F = 2$), who has befriended E, $w_{\text{dist}_{EF}} = 0$, which makes TC_E full. Finally, A ends up befriending B ($TO_B = 2$, $TC_B = 2$), $w_{\text{dist}_{AB}} = 0$, the only individual that is still available in the space and can reciprocate A's friendship.

Figure 3(d) illustrates the social network that has developed from this process. The system has reached its equilibrium since there is no additional friendship that can develop. In this case, A, F, and G, could have become friends making an open triangle considering their tie capacities are not full yet. However, since both F and G are nodes with $TO = 0$, they cannot reciprocate A's friendship initiative. A and E could have become friends if E had a higher tie capacity or if E had not befriended F. This exhibit shows how homophily prevails even in a situation where heterogeneous friendships could have developed. Four out of six (57%) dyads are between nodes of the same shape and color. Whereas only one friendship (14%) involves nodes of different colors

(but of the same shape) and another one (14%) involves nodes of the same color (but of different shapes). In the next section, we will run agent-based simulations drawn upon FIRMAN with a larger population and more trials.

Figure 3

An Example of Friendship Development Under FIRMAN

Note. Nodes can only initiate and reciprocate friendship (directed ties) within their tie outreachability (TO) range and tie capacity (TC) limit. Friendships develop among nodes, who can reciprocate each other's ties (undirected ties). (a) G ($TO_G = 0$, $TC_G = 3$) cannot reciprocate A ($TO_A = 3$, $TC_A = 3$) since G can only generate zero length of ties. (b) G can only befriend H ($TO_A = 1$, $TC_A = 2$), who fall under its tie outreachability range, ($w_{dist_{GH}} = 0$) (c) E ($TO_A = 1$, $TC_A = 1$) cannot reciprocate A's tie due to its tie capacity being full by befriendng F ($TO_A = 0$, $TC_A = 2$). (d) The system reaches its equilibrium; no more dyadic friendship can be made.

Agent-Based Simulations

Methods

Agent Populations

All agent parameters were set fixed except the parameters for tie outreachability (TO) and tie capacity (TC) that varied. For each simulation, n agents = 100, j SI = 2 with $w_1 = 1$ for SI_1 and $w_2 = 2$ for SI_2 . To generate very diverse populations, parameter p for each SI was set fixed at $p =$

.5. For SI_1 , 0 denotes round and 1 denotes square, and for SI_2 , 0 denotes white and 1 denotes black mimicking the hypothetical example (i.e., shapes and colors).

To investigate the effect of tie outreachability (TO), the parameter for TO (q) were set at .2, .3, .4, .5, .6, .7, and .8. Lower TO parameters represent populations with more individuals with either zero or one tie outreachability, while higher TO parameters represent otherwise. To investigate the effect of tie capacity (TC), the parameters for TC were set at $\mu = 4$, $\sigma^2 = 1$, and $\mu = 8$, $\sigma^2 = 1$ representing populations with relatively high and low needs for friendships, respectively. These yielded fourteen groups of scenarios as presented in Table 2. For each scenario, the simulations were repeated 500 times making total simulations equal to 7000.

Table 2

Parameters to Generate Agents

	TO $q = .2$	TO $q = .3$	TO $q = .4$	TO $q = .5$	TO $q = .6$	TO $q = .7$	TO $q = .8$
TC $\mu = 4, \sigma^2 = 1$	500 trials	500 trials	500 trials	500 trials	500 trials	500 trials	500 trials
TC $\mu = 8, \sigma^2 = 1$	500 trials	500 trials	500 trials	500 trials	500 trials	500 trials	500 trials

Note. For each trial, n agents = 100, j SI = 2 with $w_1 = 1$ for SI_1 (shape social identity) and $w_2 = 2$ for SI_2 (color social identity). For each SI, parameter p is set fixed at $p = .5$ so that it generates diverse populations with equal numbers of agents for each social identity intersection.

Simulation Procedure

For every simulation, all agents have the same goal of developing friendships within the population. The simulation begins with all agents having no friends in the system. The order in which agents get to initiate friendship first is determined by the stochastic process. The agents getting invited for a friendship cannot and will not decline unless, as explained in equation 7, their tie outreachability (TO) is shorter than the social identity distance (w_dist) between them and the agents initiating the friendship, or they have reached their maximum tie capacities (TC). All agents would continuously develop friendships until no more friendship could develop as the system has reached its equilibrium.

The model and simulations were run in R. The codes and generated data are available as supplementary materials. The pseudocode of friendship development is as follows.

```

Input: agent_population

all_possible_dyads = {choose(agent_population, 2)}
for (x, y) in random_order(all_possible_dyads):
    if w_dist(x, y) ≤ min (TO[x], TO[y])
        AND num_f(x) < TC[x] AND num_f(y) < TC[y]:
        friendship_yes = {(x, y)}
    else:
        friendship_no = {(x, y)}
return
(developed_friendships = friendship_yes)

Output: developed_friendships

```

Analysis Focus

The analysis focuses on two measures: the count of developed dyadic friendships and the number of satisfied agents in the population. As for the first measure, we use the weighted social identity distance (w_dist) to analyze to which category a friendship falls into: full homogeneous ($w_dist = 0$), shape heterogeneous but color homogeneous ($w_dist = 1$), shape homogeneous but color heterogeneous ($w_dist = 2$), or full heterogeneous ($w_dist = 3$). We then utilize chi-square tests to investigate how different the patterns of the developed friendships are from the expected ones given the proportions of agents' social identities. Should they not be significantly different, it means that the developed friendship networks do not exhibit homophily, and vice versa. This approach has been commonly used to study homophily (e.g., Thelwall, 2009).

As for the second measure, agents will be satisfied if the numbers of friends (num_f) that they make are at least *half* of their maximum tie capacity (TC) and will be not satisfied, otherwise. It should be noted that this threshold is arbitrary. Any value below one (equal to maximum TC) is reasonable. We chose half since it is in the middle, not too high nor too low.

Results

Figure 4 illustrates the agent populations generated in all simulations. As can be seen, the aggregates of all agents' social identities (SI), tie outreachability (TO), and tie capacity (TC) follow the distributions in which they have been generated from, respectively. For example, the numbers for each SI intersections (i.e., round-white, round-black, square-white, square-black) are relatively the same in all scenarios as intended. Meanwhile for both TO and TC, the proportions are similar across rows or columns in which their parameter values are held constant.

Figure 4*Agent Populations Generated in All Simulations*

Note. Top graph: The numbers of agents for each social identity intersections are relatively equal in all scenarios representing a very diverse population. Middle graph: As the tie outreachability (TO) parameter increases, the number of agents with TO = 0 decreases, and vice versa. Bottom graph: Most agents have tie capacity (TC) around means (TC = 4 for the first row, TC = 8 for the second row).

Figure 5 presents the developed dyadic friendships in all simulations. It should be noted that this graph does not represent individuals' friends (alters) as in ego networks. Instead, it visualizes pairs of friendships as in network level. As can be seen, fully homogeneous friendships dominate the simulation results in almost all scenarios. The exception is for scenarios with TO parameter $q = .8$, in which the proportions of fully homogeneous friendship ($w_dist = 0$) are almost the same as the shape heterogeneous but color homogeneous friendships ($w_dist = 1$). Furthermore, the graph also shows that doubling the numbers of TC does not change the proportion of friendships. It just increases the numbers of dyadic friendship without changing the (homogeneity/heterogeneity) patterns.

Figure 5

Developed Friendships in All Simulations

Note. Horizontal comparison: Even though increasing tie outreachability (TO) parameter (fewer agents with $TO = 0$) leads to more heterogeneous friendships in the populations, most friendships are still between agents with the same color and shape ($w_dist = 0$) in all scenarios. Vertical comparison: Doubling the tie capacity (TC) parameter increases numbers of dyadic friendships in the population but does not change their (homogeneity/heterogeneity) patterns.

To investigate whether the developed friendship exhibit homophily in a statistical sense, we performed a chi-square test for each simulation as planned. Figure 6 presents the chi-square values. A higher chi-square value means a higher gap between the proportion of developed homogeneous friendships out of the simulations and the expected dyadic friendships given the proportion of agents' social identity intersections (that should favor the development of heterogeneous friendships). As can be seen, as the TO parameter increases, the chi-square value decreases, whereas increasing TC makes chi-square values increase as well. We do not show the p-values in this manuscript since they are below the common significant threshold ($p < .001$) in all simulations highlighting that all developed friendships are above and beyond the expected friendships given the proportions of SI intersections.

Figure 6

Chi-Square Values of Developed Friendships in All Simulations

Note. Higher chi-square value represents higher gap between the proportion of developed homogenous friendships out of the simulations and the expected dyadic friendships. The proportion of agents' social identity intersections should favor heterogeneous friendships. All chi-square values are significantly different from zero ($p < .001$).

Finally, Figure 7 shows the number of satisfied agents in all simulations. As can be seen, only a miniscule number of agents in all scenarios are not satisfied with the numbers of their

friends. It means that in all scenarios, regardless of their tie outreachability level, almost all agents have friends at least half of their tie capacity.

Figure 7

Numbers of Satisfied Agents in All Simulations

Note. Only a miniscule number of agents in all scenarios are not satisfied with the numbers of their friends ($\text{num}_f < .5 * \text{TC}$).

Discussion

This paper has explained through a formal theoretical framework, FIRMAN, and a series of agent-based simulations, why homogenous friendships persist even in very diverse populations, where the probability to develop homophilous networks is *not* greater than a mere chance. In doing so, it relies on two micro mechanisms and their interactions: whether both individuals in a dyad can accept and tolerate the differences, if any, represented by tie outreachability (TO) and whether both are still available to commit for a new friendship represented by tie capacity (TC). This explanation holds in various scenarios as represented by various parameters of TO and TC, and is robust in multiple replications.

Contrary to past studies (e.g., Avin et al., 2020; Currarini et al., 2010; Wimmer & Lewis, 2010; Zeng & Xie, 2008) that tried to disentangle homophily to find its single root cause, this study observes the interactions of antecedents that may catalyze homophily on the one hand (e.g., more agents with lower tie outreachability) and hamper it on the other hand (i.e., diverse populations,

more agents with higher tie outreachability). Surprisingly, their interactions, still, disproportionately favor homophily. Even in a population where the proportion of agents who can only develop and maintain fully homogenous friendships are less than 15 %, most developed friendships (i.e., > 50% of total friendships) remain between agents who share multiple similarities. This number gets higher should we focus on one social identity at a time (in this framework: shape only or color only) as commonly used in past homophily studies (e.g., McPherson et al., 2001).

The simulation results further reveal that many agents involved in homophilous networks are not because of their preferences nor their inability to maintain heterogenous friendships, but rather due to the systemic constraints emerging from the interaction of agents with various level of TO and TC. As illustrated in the hypothetical example, friendship development does not depend only on agents' decision in the dyad, but also depend on the other agents' actions, even when they do not directly interact. For instance, round-white agents' decisions to develop friendship with the same group reduces their TC so that it cumulatively lessens other group agents' opportunities to develop fully heterogenous friendships.

Indeed, it is then tempting to say that TC is as important as TO in favoring homophily and hampering heterophily as past studies have argued, that having enough friends prevent people to develop interethnic friendships (Pfister et al., 2020). However, the simulations have shown that doubling the number of TC increases both numbers of homogeneous and heterogeneous friendships, making their proportions remain the same at the population level. Meanwhile, increasing TO while holding TC constant changes the proportions albeit the networks still exhibit homophily principle. Regardless, these findings indicate that TO is more important than TC in enabling heterogeneous friendships.

This study has also pointed out another reason why homophily persist: most agents are satisfied with their current friendships. Regardless of the deleterious effects associated with homogeneous friendships and homophilous networks (Karimi et al., 2018), most people may not see homophily as a problem. It is the case since their goal of developing friendships is to fulfill their social needs (Hojjat & Moyer, 2017). Hence, if people can maintain reciprocated friendships, regardless their SI, their social need will be satisfied. Moreover, developing homogeneous friendships are mathematically cheaper as the agents need to generate shorter length of ties ($LT =$

0) compared to heterogeneous friendships that require longer ties to connect ($LT > 1$). Of course, it would be a very different story should the people's goal in developing friendship is to develop and maintain heterogeneous friendships.

Limitations

Apart from its strengths in examining homophily involving multiple dimensions and the interactions of its antecedents in various scenarios that are otherwise difficult to be researched with human populations, there are, still, some limitations in this study. First, as it is common with simulation research, this study may have oversimplified the process of friendship development. For instance, in real life, people may not always reciprocate friendship initiative in the first attempt. There are many other considerations such as how many times they interact and how many invitations they receive from other people, something that is not taken into consideration in this study, yet. Second, this study focuses on social network level homophily. It has yet to examine how the interactions of TO and TC affect individuals as measured at ego network level. It can be the case they affect social network and ego network differently.

Future Directions

This study is just the beginning for FIRMAN in shedding more light regarding homophily. There are, still, many possible scenarios and questions regarding homophily waiting to be explored. For example, what will happen if the numbers of social identity increase? To what extent do the majority and minority proportions affect homogeneous friendships? Considering homophily with stochastic friendship algorithm is robust, under which algorithm will heterophily flourish? We suggest future studies to pick up these questions, replicate and extend our simulations with different measures.

Acknowledgement

We thank Todd L. Pittinsky, Jason J. Jones, and Gunnar Dressler for their helpful feedback. This study received no funding. The preliminary results of this study have been presented in Networks 2021 conference.

References

- Almaatouq, A., Radaelli, L., Pentland, A., & Shmueli, E. (2016). The role of reciprocity and directionality of friendship ties in promoting behavioral change. In X. K. D. Reitter, & N. Osgood (Eds.), *Social, Cultural, and Behavioral Modeling* (pp. 33–41). Springer, Cham. <https://doi.org/10.1007/978-3-319-39931-7>
- Anderson, A., Goel, S., Huber, G., Malhotra, N., & Watts, D. J. (2014). Political ideology and racial preferences in online dating. *Sociological Science*, 1(February), 28–40. <https://doi.org/10.15195/v1.a3>
- Asikainen, A., Iñiguez, G., Ureña-Carrión, J., Kaski, K., & Kivelä, M. (2020). Cumulative effects of triadic closure and homophily in social networks. *Science Advances*, 6(19), eaax7310. <https://doi.org/10.1126/sciadv.aax7310>
- Bahns, A. J., Springer, L. S., & The, C. (2015). Fostering diverse friendships: The role of beliefs about the value of diversity. *Group Processes and Intergroup Relations*, 18(4), 475–488. <https://doi.org/10.1177/1368430214566893>
- Barwick, C. (2017). Are immigrants really lacking social networking skills? The crucial role of reciprocity in building ethnically diverse networks. *Sociology*, 51(2), 410–428. <https://doi.org/10.1177/0038038515596896>
- Boda, Z., Elmer, T., Vörös, A., & Stadtfeld, C. (2020). Short-term and long-term effects of a social network intervention on friendships among university students. *Scientific Reports*, 20, 2889. <https://doi.org/10.1038/s41598-020-59594-z>
- Bogardus, E. S. (1947). Measurement of personal-group relations. *Sociometry*, 10(4), 306–311. <https://www.jstor.org/stable/2785570>
- Bratter, J. L., & King, R. B. (2008). “But will it last?": Marital instability among interracial and same-race couples. *Family Relations*, 57(2), 160–171. <https://doi.org/10.1111/j.1741-3729.2008.00491.x>
- Campigotto, N., Rapallini, C., & Rustichini, A. (2021). School friendship networks, homophily and multiculturalism: evidence from European countries. *Journal of Population Economics*. <https://doi.org/10.1007/s00148-020-00819-w>
- Christakis, N. A., & Fowler, J. H. (2014). Friendship and natural selection. *Proc Natl Acad Sci U S A*, 2014(suppl 3), 10796–10801. <https://doi.org/10.1073/pnas.1400825111/-/DCSupplemental>. www.pnas.org/cgi/doi/10.1073/pnas.1400825111

- Currarini, S., Jackson, M. O., & Pin, P. (2010). Identifying the roles of race-based choice and chance in high school friendship network formation. *Proceedings of the National Academy of Sciences of the United States of America*, 107(11), 4857–4861. <https://doi.org/10.1073/pnas.0911793107>
- Dehghani, M., Johnson, K., Hoover, J., Sagi, E., Garten, J., Parmar, N. J., Vaisey, S., Iliev, R., Graham, J., Dehghani, M., Johnson, K., Hoover, J., Garten, J., Parmar, N. J., Vaisey, S., & Graham, J. (2016). Purity homophily in social networks. *Journal of Experimental Psychology: General*, 145(3), 366–375. <https://doi.org/10.1037/xge0000139>
- Dunbar, R. I. M. (2018). The anatomy of friendship. *Trends in Cognitive Sciences*, 22(1), 32–51. <https://doi.org/10.1016/j.tics.2017.10.004>
- Fischer, M. J. (2011). Interracial contact and changes in the racial attitudes of white college students. *Social Psychology of Education*, 14(4), 547–574. <https://doi.org/10.1007/s11218-011-9161-3>
- French, D. C., Purwono, U., & Rodkin, P. C. (2012). Religiosity of Adolescents and their Friends and Network Associates: Homophily and Associations With Antisocial Behavior. *Journal of Research on Adolescence*, 22(2), 326–332. <https://doi.org/10.1111/j.1532-7795.2012.00778.x>
- Granovetter, M. S. (1973). The strength of weak ties. *American Journal of Sociology*, 78(6), 1360–1380.
- Harrigan, N., & Yap, J. (2017). Avoidance in negative ties: Inhibiting closure, reciprocity, and homophily. *Social Networks*, 48, 126–141. <https://doi.org/10.1016/j.socnet.2016.07.003>
- Heaton, T. B. (2002). Factors contributing to increasing marital stability in the United States. *Journal of Family Issues*, 23(3), 392–409. <https://doi.org/10.1177/0192513X02023003004>
- Hofstra, B., Corten, R., van Tubergen, F., & Ellison, N. B. (2017). Sources of Segregation in Social Networks: A Novel Approach Using Facebook. In *American Sociological Review* (Vol. 82, Issue 3). <https://doi.org/10.1177/0003122417705656>
- Hogg, M. A. (2016). Social identity theory. In S. McKeown, R. Haji, & N. Ferguson (Eds.), *Understanding peace and conflict through social identity theory* (pp. 3–17). Springer, Cham. https://doi.org/10.1007/978-3-319-29869-6_1
- Hogg, M. A., & Turner, J. C. (1987). Intergroup behaviour, self-stereotyping and the salience of social categories. *British Journal of Social Psychology*, 26(4), 325–340. <https://doi.org/10.1111/j.2044-8309.1987.tb00795.x>

- Hojjat, M., & Moyer, A. (Eds.). (2017). *The psychology of friendship*. Oxford University Press.
- Ilmarinen, V. J., Lönnqvist, J. E., & Paunonen, S. (2016). Similarity-attraction effects in friendship formation: Honest platoon-mates prefer each other but dishonest do not. *Personality and Individual Differences*, 92, 153–158. <https://doi.org/10.1016/j.paid.2015.12.040>
- Ilmarinen, V. J., Vainikainen, M. P., Verkasalo, M. J., & Lönnqvist, J. E. (2017). Homophilous Friendship Assortment Based on Personality Traits and Cognitive Ability in Middle Childhood: The Moderating Effect of Peer Network Size. *European Journal of Personality*, 31(3), 208–219. <https://doi.org/10.1002/per.2095>
- Iyengar, S., & Westwood, S. J. (2015). Fear and loathing across party lines: New evidence on group polarization. *American Journal of Political Science*, 59(3), 690–707. <https://doi.org/10.1111/ajps.12152>
- Jackson, C. M., Jackson, J. C., Bilkey, D., Jong, J., & Halberstadt, J. (2018). The dynamic emergence of minimal groups. *Group Processes & Intergroup Relations*, 1–9. <https://doi.org/10.1177/1368430218802636>
- Kao, G., & Joyner, K. (2004). Do race and ethnicity matter among friends? Activities among interracial, interethnic, and intraethnic adolescent friends. *Sociological Quarterly*, 45(3), 557–573. <https://doi.org/10.1111/j.1533-8525.2004.tb02303.x>
- Karimi, F., Génois, M., Wagner, C., Singer, P., & Strohmaier, M. (2018). Homophily influences ranking of minorities in social networks. *Scientific reports*, 8(1), 1–12.
- Kossinets, G., & Watts, D. J. (2009). Origins of homophily in an evolving social network. *American Journal of Sociology*, 115(2), 405–450. <https://doi.org/10.1086/599247>
- Laniado, D., Volkovich, Y., Kappler, K., & Kaltenbrunner, A. (2016). Gender homophily in online dyadic and triadic relationships. *EPJ Data Science*, 5(1). <https://doi.org/10.1140/epjds/s13688-016-0080-6>
- Lawrence, B. S., & Shah, N. P. (2020). Homophily: Measures and meaning. *Academy of Management Annals*, 14(2), 513–597. <https://doi.org/https://doi.org/10.5465/annals.2018.0147>
- Lazarsfeld, P. F., & Merton, R. K. (1954). Friendship as social process: A substantive and methodological analysis. In M. Berger, T. Abel, & C. H. Page (Eds.), *Freedom and control in modern society* (pp. 18–66). Octagon Books.

- Leszczensky, L., & Pink, S. (2017). Intra-and inter-group friendship choices of christian, muslim, and non-religious youth in Germany. *European Sociological Review*, 33(1), 72–83. <https://doi.org/10.1093/esr/jcw049>
- Mazur, E., & Richards, L. (2011). Adolescents' and emerging adults' social networking online: Homophily or diversity? *Journal of Applied Developmental Psychology*, 32(4), 180–188. <https://doi.org/10.1016/j.appdev.2011.03.001>
- McPherson, J. M., & Ranger-Moore, J. R. (2017). Evolution on a dancing landscape. *Social Forces*, 70(1), 19–42. <http://links.jstor.org/sici?sici=0037-7732%28199109%2970%3A1%3C19%3AEOADL0%3E2.O.CO%3B2-8>
- McPherson, J. M., & Smith-Lovin, L. (1987). Homophily in voluntary organizations: Status distance and the composition of face-to-face groups. *American Sociological Review*, 52(3), 370–379.
- McPherson, M. (2004). A Blau space primer: Prolegomenon to an ecology of affiliation. *Industrial and Corporate Change*, 13(1), 263–280. <https://doi.org/10.1093/icc/13.1.263>
- McPherson, M., Smith-Lovin, L., & Cook, J. M. (2001). Birds of a feather: Homophily in social networks. *Annual Review of Sociology*, 27(1), 415–444. <https://doi.org/10.1146/annurev.soc.27.1.415>
- Mollica, K. A., Gray, B., & Treviño, L. K. (2003). Racial homophily and its persistence in newcomers' social networks. *Organization Science*, 14(2), 123–136. <https://doi.org/10.1287/orsc.14.2.123.14994>
- Pfister, M., Wölfer, R., & Hewstone, M. (2020). Contact capacity and its effect on intergroup relations. *Social Psychological and Personality Science*, 11(1), 7–15. <https://doi.org/10.1177/1948550619837004>
- Poteat, V. P., Mereish, E. H., Liu, M. L., & Nam, J. S. (2011). Can friendships be bipartisan? The effects of political ideology on peer relationships. *Group Processes and Intergroup Relations*, 14(6), 819–834. <https://doi.org/10.1177/1368430211401048>
- Ramiah, A. Al, Schmid, K., Hewstone, M., & Floe, C. (2015). Why are all the White (Asian) kids sitting together in the cafeteria? Resegregation and the role of intergroup attributions and norms. *British Journal of Social Psychology*, 54(1), 100–124. <https://doi.org/10.1111/bjso.12064>
- Rude, J., & Herda, D. (2010). Best friends forever? Race and the stability of adolescent friendships. *Social Forces*, 89(2), 585–607. <https://doi.org/10.1353/sof.2010.0059>

- Simpkins, S. D., Schaefer, D. R., Price, C. D., & Vest, A. E. (2013). Adolescent friendships, BMI, and physical activity: Untangling selection and influence through longitudinal social network analysis. *Journal of Research on Adolescence*, 23(3), 537–549. <https://doi.org/10.1111/j.1532-7795.2012.00836.x>
- Smith, S., McFarland, D. A., Tubergen, F. Van, & Maas, I. (2016). Ethnic composition and friendship segregation: Differential effects for adolescent natives and immigrants. *American Journal of Sociology*. <https://doi.org/10.1086/684032>
- Stark, T. H., & Flache, A. (2012). The double edge of common interest: Ethnic segregation as an unintended byproduct of opinion homophily. *Sociology of Education*, 85(2), 179–199. <https://doi.org/10.1177/0038040711427314>
- Tajfel, H., & Turner, J. (1979). An integrative theory of intergroup conflict. In *Organizational identity: A reader* (pp. 33–47).
- Thelwall, M. (2009). Homophily in MySpace. *Journal of the American Society for Information Science and Technology*, 60(2), 219–231. <https://doi.org/10.1002/asi.20978>
- Turner, J. C., Oakes, P. J., Haslam, S. A., & McGarty, C. (1994). Self and collective: Cognition and social context. *Personality and Social Psychology Bulletin*, 20(5), 454–463. <https://doi.org/10.1177/0146167294205002>
- van Zomeren, M., Postmes, T., & Spears, R. (2008). Toward an integrative social identity model of collective action: A quantitative research synthesis of three socio-psychological perspectives. *Psychological Bulletin*, 134(4), 504–535. <https://doi.org/10.1037/0033-2909.134.4.504>
- Wark, C., & Galliher, J. F. (2007). Emory Bogardus and the origins of the social distance scale. *American Sociologist*, 38(4), 383–395. <https://doi.org/10.1007/s12108-007-9023-9>
- Wimmer, A., & Lewis, K. (2010). Beyond and below racial homophily: ERG models of a friendship network documented on Facebook. *American Journal of Sociology*, 116(2), 583–642. <https://doi.org/10.1086/653658>
- Wölfer, R., & Hewstone, M. (2018). What buffers ethnic homophily? Explaining the development of outgroup contact in adolescence. *Developmental Psychology*, 54(8), 1507–1518. <https://doi.org/10.1037/dev0000547>
- Yu, S., & Xie, Y. (2017). Preference effects on friendship choice: Evidence from an online field experiment. *Social Science Research*, 66, 201–210. <https://doi.org/10.1016/j.ssresearch.2017.01.006>