

# Segregated in Social Space

## The Spatial Structure of Acquaintanceship Networks\*

Barum Park<sup>†</sup>

### Abstract

With deepening cleavages on several social dimensions, social fragmentation has become a major concern across the social sciences. This paper proposes a spatial approach to study the segregation pattern of acquaintanceship ties across multiple social dimensions simultaneously. A Bayesian unfolding model is developed and fitted to the 2006 General Social Survey. Results suggests that the segregation pattern of reported acquaintanceship ties reflect consolidated socioeconomic inequalities. Further, among the thirteen analyzed social groups, gay and lesbian people were the least segregated group in 2006, implying that individuals with very different network compositions had similar probabilities to know someone who is gay or lesbian. Lastly, contradicting previous findings that ideology and religiosity are segregating acquaintanceship networks to an extent that rivals race, it is found that race stands out as the dominant dimension that shapes the distribution of these relationships.

Keywords: social space, network segregation, social distance, acquaintanceship ties

---

\*I am grateful to Delia Baldassarri, Siwei Cheng, Ned Crowley, Paul DiMaggio, Mike Hout, Harris Hyun-Soo Kim, Byungkyu Lee, Jeff Manza, John Levi Martin, Oscar Stuhler, and Larry Wu for their invaluable comments. This paper was presented in the Inequality Workshop at New York University, the Annual Conference of the Association of Korean Sociologists in America in 2018, and the American Sociological Association Annual Meeting in 2018. All remaining errors are my own. Direct correspondence to Barum Park, 384 Uris Hall, Department of Sociology, Cornell University, Ithaca, NY, 14853 (b.park@cornell.edu).

<sup>†</sup>Assistant Professor, Department of Sociology, Cornell University.

*In society, we not only live together, but at the same time we live apart, and human relations can always be reckoned, with more or less accuracy, in terms of distance*

Robert E. Park

*The space that I know through my senses, where I am at the center and where everything is arranged in relation to me, could not be the space as a whole, which contains all the individual spaces and in which, moreover, those individual spaces are coordinated in relation to impersonal reference points common to all individuals*

Emilie Durkheim

## **Introduction**

It is well-known that close relationships tend to generate segregated clusters of close-knit networks (Davis, 1967, 1970; Moody, 2001; Newman and Park, 2003; Martin, 2009; Rivera et al., 2010). Observing these structural tendencies, sociologists have placed much theoretical emphasis on the role of “weak ties” in binding otherwise isolated segments of society together (Granovetter, 1973; Fararo, 1992; Martin, 2009). Despite the central position of social integration in the discipline of sociology, however, it is surprising to find that most empirical research on the population distribution of social ties in the contemporary U.S. remains focused on relatively close relationships. Most of our knowledge is confined to marriage patterns, friendships, or relations of discussing important matters and politics (McPherson et al., 2001; Huckfeldt et al., 2004; Rivera et al., 2010; Schwartz, 2013) while much less is known about the distribution of the wider networks of the public (DiPrete et al., 2011; Hofstra et al., 2017). This lack of knowledge is unfortunate, particularly in an era in which social fragmentation has become a major concern among social scientists and the general public.

Although many claims regarding social and political fragmentation are exaggerated (Fiorena and Abrams, 2008; Fischer and Mattson, 2009), they are not completely unfounded. In light of the well-established finding that social relationships tend to be formed between individuals with similar social characteristics (McPherson et al., 2001; Smith et al., 2014), the increased consolidation between socioeconomic background, political views, cultural lifestyles, and geography that developed over the last decades might create barriers across which social ties are

less likely to form (Blau, 1977; Skvoretz, 1983; Centola, 2015).<sup>1</sup> Indeed, while race and socioeconomic status are regarded as the strongest basis of homophily (McPherson et al., 2001), political identities are increasingly influencing whom individuals date and befriend, which, in part, can be traced back to their increased consolidation with other social attributes (Klofstad et al., 2013; Iyengar et al., 2019; Mason, 2018; cf. Huber and Malhotra, 2017). Yet again, we find that most of these findings are based on close relationships that fail to capture social ties that are less demanding of time and emotional investment and, thus, less influenced by homophilous biases (Rivera et al., 2010).

In an important contribution, DiPrete et al. (2011) introduced new data and methods to analyze the distribution of acquaintanceship ties from randomly sampled ego-networks. The approach was specifically designed to capture the “structure of complete social networks—including the weak ties as well as the strong ones.” (p. 1240.) Confirming worries about social fragmentation, their analysis showed that network segregation of acquaintanceship ties “based on religiosity, political ideology, family behaviors, and socioeconomic standing are high and in some cases rival racial segregation in their intensity” (p. 1236). However, by analyzing single social groups<sup>2</sup> in isolation, the analytical approach adopted by DiPrete and colleagues overlooked a unique aspect of networks that is crucial for the study of segregation: namely, that multiple groups co-appear simultaneously in each ego-network and that these groups are indirectly linked to one another through their common connection to the sampled ego.

In this paper, I push the pioneering work of DiPrete et al. (2011) one step further by adopting an analytical approach that makes full use of the pattern in which multiple groups intersect in the networks of individuals. I argue that this approach overcomes two important limitations of previous approaches that study network segregation from sampled ego-centric network data.

---

<sup>1</sup>On the tightened consolidation between education and economic outcomes, see Fischer and Hout (2006: Ch. 6); consolidation between political ideology and partisan identities are reported in DiMaggio et al. (1996), Baldassarri and Gelman (2008), and Park (2018); partisanship has also become better aligned with other social identities such as race, religion, and gender (Kaufmann, 2002; Mason, 2018) and political views have been found to be increasingly correlated with income (McCarty et al., 2006), religiosity (Hout and Fischer, 2014), lifestyles (DellaPosta et al., 2015: 1506), and geography (Bishop, 2009; but see, Mummolo and Nall, 2017); similar results on geographic clustering are also found for cultural lifestyles (Fischer and Mattson, 2009) and income (Reardon and Bischoff, 2011)

<sup>2</sup>The term “social group” is used here in a loose way, referring to individuals that share social attributes, regardless of whether members are in direct contact with one another or how much they identify themselves with the attribute. For example, individuals with the same educational background or political ideology are defined to belong to the same social group.

First, it enables the simultaneous analysis of segregation patterns between multiple groups instead of focusing on one group at a time or isolated pairs of groups and, second, it goes beyond quantifying the *extent* to which acquaintanceship networks are segregated to showing the structure of *how* they are segregated.

The main device to achieve these goals is to conceptualize individuals and groups as occupying positions in a joint “social space” in which proximity is a function of the number of ties that link individuals to members of different social groups. Since individuals maintain multiple ties to multiple groups, proximities in the space are defined in terms of the pattern by which social groups intersect in individuals’ social networks: groups that co-appear frequently in the networks of individuals will be close to each other, while groups that appear only rarely together will be distant. In this way, the entire information contained in each individual’s network composition is used to calibrate the positions of groups, which includes their distances between each other as well as their distances to all individuals in the space. I demonstrate how this conceptualization leads to a space that is defined in terms of structurally equivalent positions of individuals and groups vis-a-vis each other. As the positions are derived from the entire distribution of social ties that link all individuals to groups, the space offers a holistic view on the segregation structure of these networks. Further, it highlights structural aspects of the distribution of social ties that are not directly observable when single or pairs of groups are analyzed in isolation and only arise through the interdependence between group- and individual-positions in the space.

The idea to represent social structures as topological spaces is, of course, not new. Indeed, it is almost as old as the sociological discipline itself (e.g., Simmel, 1908: esp. Ch. IX, VI; Park, 1915; Sorokin, [1927]1959) and can be found throughout the literature on social networks (Laumann and Guttman, 1966; Laumann and Pappi, 1976; Winship, 1977; Burt, 1990; Freeman and Webster, 1994; Bottero and Prandy, 2003; Martin, 2009; Smith, 2017). This is not surprising as social structures are often, if not always, conceptualized in terms of social positions which are “close” or “far” and lie “below” or “above” one another (Laumann and Pappi, 1976; Bourdieu, 1989). The reason for relying on spatial models lies, however, not only in their intuitive appeal; just as each node in a network gains meaning through its relations to other nodes,

so is the meaning of a point in a space determined through its relations to other points. Thus, both spatial metaphors and networks preserve a fundamental notion of social structures: their *relative* nature. Preserving the relational aspect of networks is particularly important to study network segregation. Groups can never be *just* segregated but are always segregated *from* other groups, analogous to the fact that a point in space can never be just far but has to be far from another point. As social distances tap naturally into the notion of network segregation, I use the collection of these distances as well as their interrelations to analyze spaces that reflect the segregation patterns across multiple groups simultaneously.

On the methodological side, conceptualizing social ties as being formed according to proximities in a space has the benefit of being easily translated into a formal statistical model that can be used to analyze empirical data. After discussing how the co-appearance pattern of groups within individuals' social networks can be used to define distances in social space, an unfolding model is formulated that incorporates these intuitions. The model is, thereafter, fitted to a battery of questions in the 2006 General Social Survey (GSS) that asks about the number of acquaintanceship ties to thirteen groups that were chosen to represent the major social cleavages in the contemporary U.S. (DiPrete et al., 2011).

Results of the analysis show that the estimated social space, and hence network segregation patterns of reported acquaintanceship ties, reflect consolidated socioeconomic inequalities. Yet, contrary to claims that ideology and religiosity are becoming major lines of segregation and conflict in the U.S. (Hunter, 1991; Bishop, 2009; Abramowitz, 2010; DiPrete et al., 2011; Iyengar and Westwood, 2015), I find that these dimensions of differentiation were of relatively minor importance when it comes to the distribution of acquaintanceship ties in 2006. Instead, race was the most salient social dimension that segregated these relationships. Lastly, the group that is positioned closest to the center of the social space are gay and lesbian people. This implies that individuals with very different network compositions were similarly likely to report that they know someone who is gay. In other words, at the same time as political polarization on gay rights issues was reaching its apex of divisiveness in contemporary U.S. history (Hetherington, 2009), gay and lesbian people were surprisingly well-integrated in the reported acquaintanceship networks of U.S. citizens.

The paper unfolds as follows. In the next section, I discuss how social distances between social groups can be defined in terms of their co-appearance pattern in randomly sampled ego-networks. Thereafter, a model of social space is presented together with a discussion of how positions in the space can be used to study network segregation. After discussing the limitations of the approach, the model is fitted to the 2006 GSS. The paper concludes with a discussion of the findings and their implications.

## **Networks and Spaces**

Network scholars face a dilemma when studying large populations. While face-to-face interactions are deemed essential for the integration of societies (Blau, 1977; Fararo, 1992; see also, Simmel, 1908: Ch. I, VII), it is often impossible to observe even a small subset of the associations between members of large collectivities such as cities or nations. Hence, researchers are forced to confine their study to relatively small populations with well-defined boundaries, where the relationship between all individuals can be assessed, or rely on randomly sampled “ego-centric” networks, and the group-belongings of the sampled ego and alters, when studying larger ones. Yet, the very characteristics of random samples that enable researchers to make valid statistical claims about large populations come with the caveat that the sampled individuals are, by design, not directly connected to one another. Out of this reason, research that utilizes representative samples to study social distances and network segregation focus on the differences in social attributes between ego-alter pairs, measures of heterogeneity within sampled ego-networks, or the comparison of the observed distribution of ties to random mixing scenarios (Marsden, 1987, 1988; McPherson et al., 2006; DiPrete et al., 2011; Smith et al., 2014).

Research on ego-centric networks has accumulated a substantial amount of evidence that close social ties tend to be maintained between individuals that belong to the same social group (McPherson et al., 2001; Rivera et al., 2010), which reflects a structural tendency towards network segregation. Nevertheless, two important limitations remain. First, most previous studies focus on relatively “strong” ties and hard evidence regarding the segregation pattern of broader

networks in the contemporary U.S. remains scarce (DiPrete et al., 2011). The lack of evidence regarding broader networks that include weak relationships implies that exactly those types of ties that are deemed to cross group-boundaries, and thus are essential for binding different parts of society together, were not captured by these studies (Granovetter, 1973; Fararo, 1992).

A second limitation lies in the tendency to analyze single groups or single dimensions of differentiation in isolation. Associational patterns are often analyzed along one social dimension at a time—such as associations between people with different racial or educational backgrounds (e.g., Marsden, 1988; McPherson et al., 2006)—or by comparing the number of ties a group receives to what would be expected when the population mixes randomly (e.g., DiPrete et al., 2011). However, less is known about how these patterns are combined *across* several dimensions and how the between-group distances are structurally related to one another. Hence, these studies offer us only a partial picture of the association patterns in society as a whole (Wimmer and Lewis, 2010; Centola, 2015) and a key element that determines the structure of network segregation goes missing—namely, how associating with members of one group constrains or promotes the association with other groups.<sup>3</sup>

## FROM DYADS TO NETWORKS

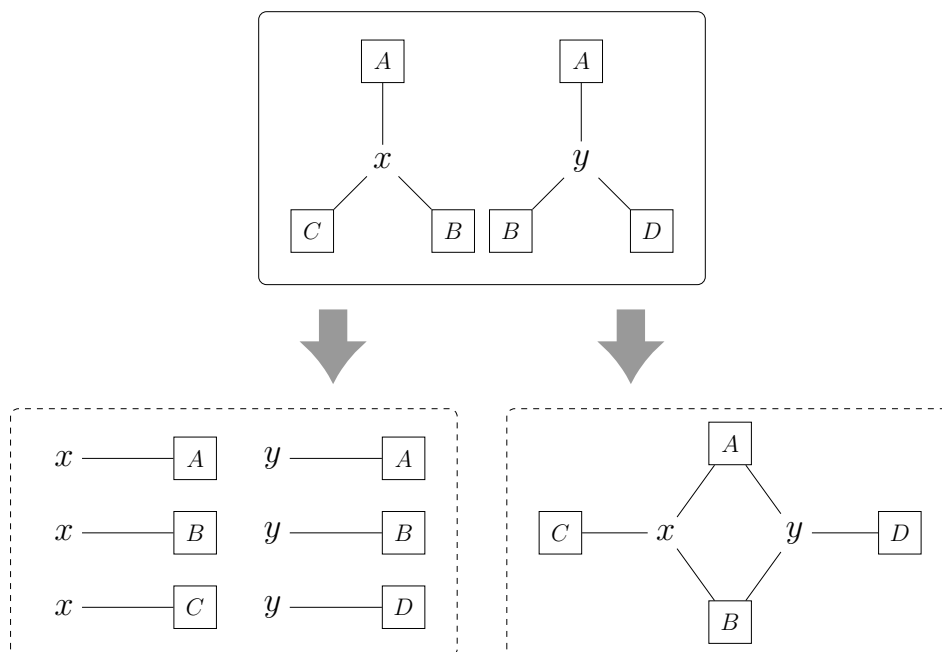
But how can we analyze network segregation patterns on multiple dimensions simultaneously? I argue that this can be achieved by switching the analytical focus towards the entire ego-centric network in which the sampled respondents are embedded and the *co-appearance pattern* of groups within them. Each ego-network connects the sampled respondent (the ego) to a number of other individuals (the alters) who, in turn, belong to multiple groups. In order to study social distances between all considered groups simultaneously, I use the fact that all groups that co-appear in the ego-network are indirectly tied to each other through the ties spawned by the ego. Hence, the focus is on all the two-paths found in the ego-network of the form “*A*–ego–*B*” that link the groups, *A* and *B*, together.

---

<sup>3</sup>In this regard, the study by Smith et al. (2014) offers an important step forward by controlling for the consolidation/intersection of several social dimensions when estimating homophilous tendencies in network formation. However, controlling for between-dimension consolidation in regression models to obtain “net-effects” of homophily is different from modeling the interdependencies between different dimensions themselves. It is the latter, namely how being connected to members of one group is related to having ties to members of multiple other groups, that is of major importance to understand the segregation pattern between multiple groups.

Figure 1 illustrates the difference between previous approaches (left) and the approach used here (right) in analyzing ego-centric network data. In most previous approaches to study network segregation or social distances, the ego-network is first disaggregated into isolated dyads. In the case of DiPrete et al. (2011), the distribution of ties that a group receives is thereafter compared to expectations under randomly mixing populations. In other approaches, the group-memberships of the respondents are compared to that of alters to examine social distances between the groups (e.g., Marsden, 1987; McPherson et al., 2006; Smith et al., 2014). By disaggregating the ego-network into dyadic relationships, however, the tie between  $x$  and  $A$  is treated as being independent of the tie between  $x$  to  $B$ , which ignores the fact that  $A$  and  $B$  appear together in same individuals' ego-network and are, thus, indirectly linked.

Figure 1: Illustration of Approach to Analyze Ego-centric Network Data



Notes: Respondents are represented by lowercase letters, while groups are represented by uppercase letters in squares.

In this paper, the simultaneous co-appearance of groups within individuals' ego-networks is the very starting point in defining social distances between them. Individuals are conceptualized to be close to the groups to which they have many connections, while the groups are close if they co-appear frequently in individuals' networks. To make sense of this definition



of social distance, consider an individual, call her Anna, who moves to Seattle after spending her life in New York City. The relations to her New Yorker friends are what defines how socially close Anna is to New York City; similarly, by befriending new friends in Seattle, she'll become closer to her new hometown. Yet, the same new connections that bring Anna socially closer to her new home, will bring the groups Seattle and New York closer to each other as well. In other words, by maintaining ties to two groups simultaneously—here Seattle and New York City—each individual is reducing the social chasm between them. More generally, by befriending members of a group, the individual not only becomes closer to the group herself but, simultaneously, brings the group closer to *all other groups* that appear in her network. The avoidance of certain groups, in contrast, pushes the focal group away from all other groups to which the individual is connected.

The lower-right plot of Figure 1 shows a simplified version of how these intuitions play out in the case of two individuals and their network compositions. Groups  $A$  and  $B$  appear simultaneously in the networks of both  $x$  and  $y$  and are, thus, placed closer to each other than  $C$  and  $D$ , which appear only in either  $x$ 's or  $y$ 's networks but not in both. With only two individuals, it is impossible to know how far the groups will be apart; we know only that  $A$  and  $B$  are closer to each other than  $C$  and  $D$ . Yet, when averaged over many ego-networks, it is possible to calibrate the positions of the groups from the set of individuals' network compositions. In this way, the full set of connections that an individual holds can be used to study distances between not only isolated pairs of groups but between multiple groups simultaneously.

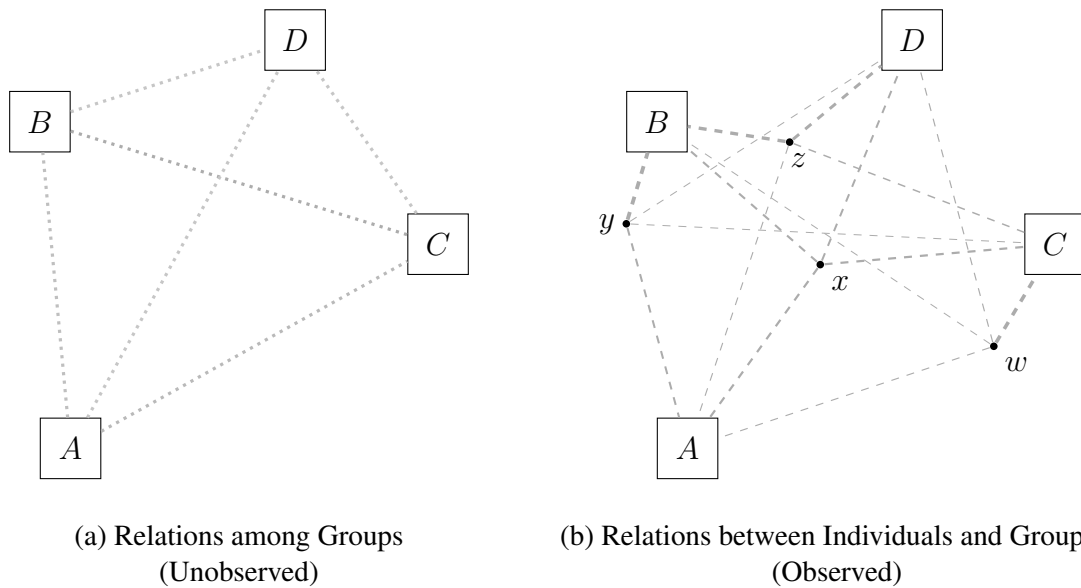
## FROM NETWORKS TO SPACES

The way in which social distances are defined naturally informs the metric of the social space corresponding to it. If individuals are treated as being close to groups to which they have many connections while groups that co-appear frequently together in individuals' networks are close to each other as well, it should be the case that two groups that are close to many individuals cannot be too distant in social space.<sup>4</sup> Figure 2 shows a schematic representation of a space that

---

<sup>4</sup>The most natural way to formalize this notion is by asserting that distances in the social space satisfy the triangle inequality. Notice that the triangle inequality rules out many alternative ways in which social distances

Figure 2: Positions of Individuals and Groups in Social Space, Schematic Representation



Notes: Individuals are represented by points and lowercase letters. Groups are represented by uppercase letters in squares. The width of the lines that connect individuals to groups is proportional to the number of ties that individuals have to the groups.

satisfies this condition. The positions of individuals are denoted by lower-case letters  $w, x, y, z$ , and groups by the capital letters  $A, B, C$ , and  $D$ . The plot on the left shows the set of distances between the groups in social space. These distances are not observed and have to be inferred from the data, similarly to the distances between the individual positions. The plot on the right shows, on the other hand, what is observed: a set of sampled individuals and the number of ties they hold with members of different groups.

The distance of an individual to a groups is determined by the relative prevalence of the group in the individual’s ego-network, so that an individual with many ties to members of a certain group is placed closer to it. For example, the position of  $x$  would imply that she holds a similar number of ties to members of all groups, while  $z$  will have more ties to  $B$  and  $D$  than to  $A$  and  $C$ .<sup>5</sup> As each point in this space is defined in relation to all other points, the position

could be defined. This includes the “minimum ultrametric” (Watts et al., 2002) which has gained considerable attention from sociologists to model the boundary-crossing nature of weak ties (Martin, 2009: Ch. 2, Centola, 2015). A short justification of ruling out the minimum ultrametric is laid out in the supplementary materials.

<sup>5</sup>For the sake of illustration, it is assumed that the total number of ties that all individuals hold are the same and that the population shares of each group are equal to one another. In the model developed below, however, the number ties that each individual holds as well as the population shares of the groups will be adjusted for. Also, the

of each individual is determined not by her relations to each social group in isolation but by the whole network of ties in which she is embedded in. Thus, two individuals will occupy the same point in social space only if they have the exact same number of ties to *all* groups in the space.

On the other hand, viewed from the perspective of groups, each point in the space defines a unique combination of connections to all individuals. Group *A*, for example, will receive similar numbers of ties from *w*, *x*, and *y*, as the group is approximately equally distant from these individuals. The same is true for *D* with respect to *w*, *x*, and *y*. But as *D* receives more ties from *z* than *A*, these groups are placed on opposite sides of the space. Two groups would be placed at the same point only if they receive the exact same number of ties from *all* individuals. In short, groups that occupy the same point in the space are structurally equivalent (Lorrain and White, 1971) with respect to their relations to all individuals, while individuals occupying the same point in the space are structurally equivalent with respect to their relations to all groups.

Notice that the space so constructed will reflect two key components of Peter M. Blau macrostructural theory (Blau, 1977, 1994; Blau and Schwartz, [1984]1997) despite being inherently different from what is known as the “Blau space” (McPherson and Ranger-Moore, 1991; McPherson, 2004; Brashears et al., 2017).<sup>6</sup> First, the *salience* of group attributes—i.e., the extent to which the number of in-group ties exceed chance expectations of a randomly mixing population—is reflected in the distance between individuals belonging to a group and the position of that group in social space, as a shorter distance between the two indicates the prevalence of in-group ties. The *consolidation* of group boundaries—i.e., the population correlation between group attributes—is reflected indirectly: For two groups with perfectly consolidated boundaries—i.e., if two group attributes are perfectly correlated in the population—associating with a member of one group logically implies the association with a member of the other, so that both the number of their incoming ties as well as the source of these ties are necessarily equal. Accordingly, the positions of these groups will be exactly the same. Hence, the

---

reference to a concrete actor or individual is not accurate as the point *x* represents a position that might be filled by one, many, or no actors. However, for expository reasons, I will refer to points in the space as actors.

<sup>6</sup>In the Blau space, each dimension is defined in terms of a group attribute, while the dimensions of the space presented here have no substantive meaning. Indeed, only the distances between points in the space can be interpreted meaningfully in terms of rates of interaction, so that any isometric transformation of Figure 2 would not change the interpretation of it.

consolidation of group boundaries at the population level will put an upper bound on the possible distances between any two groups in social space. Negatively correlated group-attributes would, however, not necessarily lead to more distant group-positions.<sup>7</sup>

## The Model

The intuition of the social space representation in Figure 2 is easily translated into a formal statistical model. Since the number of alters to which an individual is connected has to be a non-negative integer, I follow DiPrete et al. (2011) and start with the assumption that the number of social ties individual  $i$  has to a member in group  $j$ , denoted by  $y_{ij}^*$ , follows a Poisson distribution with rate parameter  $\mu_{ij}$ , i.e.,

$$y_{ij}^* \sim \text{Poisson}(\mu_{ij}). \quad (1)$$

As will be discussed shortly,  $y_{ij}^*$  is not directly observed in the analyzed dataset. But for now, let us assume that we can directly measure  $y_{ij}^*$ . The distribution of ties for a population that mixes randomly can be approximated by setting

$$\mu_{ij} = \gamma_i \pi_j \quad (2)$$

where  $\pi_j$  represents the population share for group  $j$ , and  $\gamma_i$  is the gregariousness or out-degree of individual  $i$ . In other words, if individuals would create their ties randomly, they would allocate their connections proportional to the size of each group in the population.

While  $\gamma_i$  can be estimated from the data by counting the number of ties each individual holds, estimation of  $\pi_j$  has to rely on external data sources. This is because the number of ties that a group receives will inevitably conflate the population share with the associational biases toward that group. For example, even if 20% of the population belongs to group  $j$ , the group will receive less than 20% of the observed ties if there are either individual or structural reasons

---

<sup>7</sup>This is because individuals can have multiple ties to different alters. For instance, even if no low-income person in the population is Asian, it would be still possible for individuals to have ties to both Asians and low-income alters, even though it is likely that negative consolidation would lead to considerably lower co-appearance rates in any real dataset.

that hinder associating with their members. In an extreme case where the dataset contains no incoming ties to group  $j$  at all,  $\pi_j$  would be estimated to be zero even if 20% of the population belongs to the group and the same share of incoming ties is expected under random mixing. Thus, letting  $\pi_j$  be the share of received ties leads to a model in which the *observed* ties are randomly rewired under the constraint that the out-degree of individuals and the in-degree of groups are fixed, not one in which individuals maintain relationships irrespective of others' group-belongings.

I use deviations from the random mixing scenario to estimate positions in social space. Let  $\theta_i = [\theta_{i1}, \dots, \theta_{iD}]^\top$  and  $\xi_j = [\xi_{j1}, \dots, \xi_{jD}]^\top$  be, respectively, the position of individual  $i$  and groups  $j$  in a  $D$ -dimensional social space, and let  $\rho(\theta_i, \xi_j)$  be the distance between them. For the current application, I represent the distances in social space by Euclidean distances:

$$\rho_{ij} = \rho(\theta_i, \xi_j) = \left( \sum_{d=1}^D (\theta_{id} - \xi_{jd})^2 \right)^{1/2} .$$

The choice of the Euclidean metric stems largely from considerations regarding interpretability, although, in principle, any proper norm could be used to define distances in the space.<sup>8</sup> Let  $\delta$  be the rate at which distances in social space are translated into interpersonal contacts. By adding these terms to the model, we obtain

$$\begin{aligned} y_{ij}^* &\sim \text{Poisson}(\mu_{ij}) \\ \mu_{ij} &= \gamma_i \pi_j \rho_{ij}^{-\delta} \end{aligned} \tag{3}$$

Formulated in this way, it becomes apparent that  $\gamma_i$  and  $\pi_j$  are nothing else than offsets which are often used in Poisson regressions. Hence, the estimated positions will reflect the biases—either individual or structural—in the distribution of social ties connecting individuals to groups, after the population distribution of the groups and the gregariousness of individuals are adjusted for.

---

<sup>8</sup>For example, a unit circle—i.e., the set of points which are of distance 1 from the origin—has the shape of a “box” or a “diamond,” respectively, in a two-dimensional space of real numbers equipped with the infinity norm or the absolute-value ( $L^1$ ) norm. Operations such as rotations or projections are even more difficult, although not impossible, to define with these metrics. Hence, the substantive meaning any graphical representation of the space would be hard to grasp, even if such a space could be estimated and might lead to a better fit.

Lastly, to account for the possibility that the distribution of ties is overdispersed, we add a dispersion term  $\phi > 0$  to the model and assume

$$\begin{aligned} y_{ij}^* &\sim \text{Negative Binomial}(\mu_{ij}, \phi) \\ \mu_{ij} &= \gamma_i \pi_j \rho_{ij}^{-\delta}, \end{aligned} \tag{4}$$

with  $E[y_{ij}^*] = \mu_{ij}$  and  $\text{Var}[y_{ij}^*] = \mu_{ij} + \frac{\mu_{ij}^2}{\phi}$ .<sup>9</sup> The assumption of this model is different from that in DiPrete et al. (2011) in that only a single dispersion parameter is estimated for all groups. Hence, we are assuming that there is a constant rate at which the distribution of ties to all groups is overdispersed, while other variations in the tie-distribution are explained by locations in social space. This completes the model.

## Data, Estimation, and Measurement

### DATA

The analyzed data come from the 2006 General Social Survey. Respondents were asked the question “how many people that you are acquainted with are X?” where X refers to thirteen social groups defined in terms of economic status, race, ideology, and religiosity, among others. This battery was originally analyzed by DiPrete et al. (2011) and measures acquaintanceship ties to “highly salient groups that define important sources of heterogeneity among Americans and potentially important sources of social cleavage”(DiPrete et al., 2011: p. 1242), which makes the data particularly adequate to study the segregation structure of acquaintanceship ties in the contemporary U.S.

In comparison to the “important matters” name generator (Burt, 1984), the analyzed battery relaxes the notorious boundary specification problem by letting respondents enumerate any finite number of ties, although in a truncated manner, to all groups, whereas the important matter

---

<sup>9</sup>This is equivalent to adding an independently distributed multiplicative error term  $\epsilon_{ij} = \mu_{ij}/(\gamma_i \pi_j \rho_{ij}^{-\delta})$  to the mean-specification of the model and assuming that  $\epsilon_{ij}$  follows a Gamma distribution with shape and rate parameter equal to  $\phi$ . Integrating  $\phi$  out will lead to the model in (4). Also, the assumption that  $E[\epsilon_{ij}] = \phi/\phi = 1$  is innocuous here due to the  $\gamma_i$ 's in the model. Adding a free parameter for the mean of  $\epsilon_{ij}$ , while setting the geometric mean of the  $\gamma_i$ 's to one would result in the same likelihood.

name generator usually asks only for five alters. Also, while what individuals consider to be an important matter is ambiguous (Bearman and Parigi, 2004), the meaning of acquaintanceship is clearly specified in the battery used here as someone “you know their name and would stop and talk at least for a moment if you ran into the person on the street or in a shopping mall.”

Respondents reported the number acquaintanceships in each group using five response categories: “0”, “1”, “2-5”, “6-10”, and “more than 10.” Due to the split design of the 2006 GSS, 672 respondents were randomly subsampled to answer questions about their acquaintanceship networks. It should be noted that the reduced samples size will increase the uncertainty in the obtained estimates, while it will not threaten the representativeness of the analyzed sample. In the analysis that follows, I exclude respondents who reported to have no ties to any of the thirteen social groups, as it is impossible to estimate their position, and those who gave less than six valid responses. This leaves a total of 8,069 acquaintanceship ties connecting 655 respondents to the thirteen groups. The groups about which respondents were asked are shown in Table 1.

## ESTIMATION

Parameters of the model are estimated using Bayesian inference. Most quantities of interest in the analysis are functions of the unobserved positions in social space. Yet, maximum likelihood estimation techniques have inherent difficulties in estimating the  $\theta_i$  and  $\xi_j$  vectors, especially their uncertainty. In the Bayesian framework, on the other hand, the posterior distribution of these positions will summarize our uncertainty in the estimates, plus any function thereof. This property will be important to make valid statistical claims about the quantities that are of interest in this paper. In addition, even our best estimates of the population shares of the social groups are subject to various sources of error. These uncertainties can be incorporated into the model by specifying informative prior distributions, rather than using fixed values, for the proportion of each group in the model. Each  $\pi_j$  is given a Normally distributed prior truncated from below at zero and from above at one. The location and scale parameters of the priors are shown in the second and third column of the Table 1.

Lastly, as the number of acquaintanceships to each group is asked using five categories, the

Table 1: Prior Means, Scale Parameters, and Data Sources of Population Shares of Analyzed Social Groups

Group	Abbreviation	$\mu_{\pi_j}$	$\sigma_{\pi_j}$	Source <sup>a</sup>	Year
Unemployed	Unemp	0.055	0.010	BLS	2006
Own second home	2nd. Home	0.061	0.010	PSID	2005
In state or federal prison	Prison	0.010	0.005	BJS	2006
Asian or Asian-American	Asian	0.043	0.010	ACS	2006
Black or African-American	Black	0.122	0.010	ACS	2006
Hispanic	Hispanic	0.148	0.010	ACS	2006
White	White	0.662	0.010	ACS	2006
Gay men or women	Gay/Les	0.014	0.020	NSFG	2006-2008
Cohabiting women	Cohab	0.022	0.010	CPS	2007
Attend religious service regularly <sup>b</sup>	Rel. Reg	0.311	0.020	GSS	2006
Attend religious service rarely/never <sup>c</sup>	Rel. Never	0.422	0.020	GSS	2006
Strongly liberal <sup>d</sup>	Lib	0.088	0.019	GSS	2006
Strongly conservative <sup>d</sup>	Cons	0.116	0.027	GSS	2006

<sup>a</sup> BLS = Bureau of Labor Statistics (U-5), PSID = Panel Study of Income Dynamics, BJS = Bureau of Justice Statistics, ACS = American Community Survey, NSFG = National Study of Family Growth as reported in Gates (2011), CPS = Current Population Survey (Annual Social and Economic Supplement), GSS = General Social Survey. The location parameters are equal to the estimated proportions of the groups' population share.

<sup>b</sup> Proportion of respondents attending religious service "nearly every week" or more.

<sup>c</sup> Proportion of respondents attending religious service "once a year" or less.

<sup>d</sup> The location parameter is set equal to the estimated midpoint between the population share of respondents identifying as "extremely liberal" and "liberal." The scale parameter is chosen such that 99% of the prior density lies within the estimated share of these groups. The prior mean and standard deviation for Strongly Conservative group are chosen analogously.

likelihood of the data follows not a straightforward Negative Binomial distribution, but rather a collapsed version of it:

$$\begin{aligned}
 p(Y|\Xi) &= \prod_{i=1}^N \prod_{j=1}^J \prod_{k=1}^K \left[ \sum_{y_{ij}^* \in I_k} \text{Negative Binomial}(y_{ij}^* | \mu_{ij}, \phi) \right]^{\mathbb{I}(y_{ij}^* \in I_k)} \\
 &= \prod_{i=1}^N \prod_{j=1}^J \prod_{k=1}^K \left[ \sum_{y_{ij}^* \in I_k} \binom{y_{ij}^* + \phi - 1}{y_{ij}^*} \left( \frac{\mu_{ij}}{\mu + \phi} \right)^{y_{ij}^*} \left( \frac{\phi}{\mu_{ij} + \phi} \right)^\phi \right]^{\mathbb{I}(y_{ij}^* \in I_k)} \quad (5)
 \end{aligned}$$

where  $Y$  is the response matrix of dimension  $N \times J$ ,  $N$  the number of respondents,  $J$  the number of groups,  $K$  the number of response categories, and  $I_k \in \{\{0\}, \{1\}, \{2, 3, 4, 5\}, \{6, 7, \dots, 10\}, \{11, 12, \dots\}\}$  the number of ties corresponding to each response category.  $\mu_{ij}$  and  $\phi$  are defined in Equation (4).



The model is thus a variant of the unfolding model (Borg and Groenen, 2005), where we know that the responses are discrete count variables, but where the responses are only partially observed. It is also similar to the latent space model (Hoff et al., 2002) applied to a two-mode network with likelihood function as specified in Equation (5) and where the groups' population shares are estimated from exogenous data sources. A similar approach was introduced by McCormick and Zheng (2015), where arc-distances on the unit hypersphere were used represent interaction patterns. Here we use Euclidean distances, since they are more intuitive and easier to represent graphically.

The model has two main advantages over other commonly used methods to analyze two-mode network data. First, when two-mode networks are projected into their one-mode counterparts (Breiger, 1974), important information regarding the underlying structure of the network is lost. For example, when the individual-to-group response matrix is projected into a group-to-group matrix, the cells of the resulting matrix will contain only the number of individuals who hold ties to two groups simultaneously, whereas information regarding who these individuals are and to which other groups they are connected to is lost. Yet, as explained in the section that follows, it is precisely this information that is crucial to examine network segregation from the data that is analyzed. Second, the proposed model improves on correspondence analysis in that the distances in the space can be directly interpreted as Euclidean distances, which is not possible for the "spaces" generated by correspondence analysis (Borgatti and Everett, 1997; Wasserman and Faust, 1994). Indeed, while the between-individual and between-group distances in the plots generated by correspondence analysis might be interpreted as approximate chi-squared distances, the individual-to-group distances are essentially "meaningless" (Greenacre and Hastie, 1987: 441).

Clearly, the likelihood in Equation (5) is not identified. In order to identify the likelihood, I use a set of hard and sign constraints that are laid out in the supplementary materials. As all parameters in the model are continuous, I use the Hamiltonian Monte Carlo (HMC) algorithm (the No-U-Turn Sampler) implemented in `Stan` to obtain samples from the posterior distribution. Two chains are run with 7,000 iterations each, where the first 3,000 iterations are used to tune the algorithm ("warm-up") and the last 4,000 iterations are sampled for posterior

inference. This results in 8,000 samples from the joint posterior distribution of the parameters. All models that are presented below showed strong signs of convergence with the estimated potential scale reduction factor ( $\hat{R}$ ) being below 1.013 for all parameters in all models. None of the post-warm-up iterations resulted in divergences. Further details about the model, prior specifications, convergence diagnostics, and the estimation procedure can be found in the supplementary materials.

## MEASURING NETWORK SEGREGATION

The joint space of individuals and groups highlights that there are two distinct ways in which we can think about network segregation. Recall that groups which co-appear frequently in the networks of individuals are placed close to each other in social space. In Figure 2 for example, where we have assumed that the groups' population shares are equal, group  $B$  and  $D$  would appear more frequently together in individuals' networks than  $D$  and  $A$ , as the distance between the former pair is smaller than that of the latter. Since the statistical model adjusts for individuals' gregariousness and the population shares of the groups, the between-group distance in social space is a straightforward measure for the *co-appearance rates* of groups in the acquaintanceship networks of the respondents after these adjustments.

However, the co-appearance rate is different from network segregation. Consider, for example, groups  $A$  and  $D$ , in Figure 2. This pair has the largest between-group distance, and is approximately equally close to the individuals  $w$ ,  $x$ , and  $y$ . Hence, while the groups co-appear rarely in these individuals' networks, the network composition of  $w$ ,  $x$ , and  $y$  will not differ much with respect to the groups  $A$  and  $D$ . The reason for their infrequent co-appearance is not because one group appears frequently in the networks of one subset of individuals and the other group frequently in a different subset, but because both appear rarely in *all* of the individuals' networks. Hence, the co-appearances rate alone, while sufficient to conceptualize social distances, cannot distinguish between isolation and segregation.

In order to go beyond the co-appearance rates and to examine which pairs of groups are segregating acquaintanceship networks—in the sense that their positions reflect varying network-compositions across different subpopulations—we have to take into account not only the dis-

tances between groups but also the distribution of individuals in the space. In this paper I define acquaintanceship networks to be *segregated* with respect to a pair of groups if some respondents tend to be close to one group and distant to the other, while the opposite is true for another set of respondents. In other words, to talk about segregation it is not sufficient that two groups do not appear together in individuals' networks; rather, they have to appear frequently in *distinct* individuals' networks (cf. Freeman, 1978). Compare, for example, the pair of groups  $\{B, C\}$  in Figure 2 with  $\{A, D\}$ . In contrast to groups  $A$  and  $D$ , which appear at similar rates in the networks of  $w$  and  $y$ , group  $B$  mainly appears in the network of  $y$  and only rarely in the network of  $w$ , while the opposite is true for group  $C$ , which frequently appears in  $w$ 's network but not in  $y$ 's. Thus, rather than the distance between the groups per se, it is the amount of variation in the distances individuals have to the groups that is essential to measure network segregation.

Let us denote the relative distance between two groups,  $j$  and  $k$ , from the perspective of individual  $i$  by  $\Delta_{jk}(i) = \rho_{ij} - \rho_{ik}$ , where  $\rho_{ij} = \rho(\theta_i, \xi_j)$  is the distance from individual  $i$  to group  $j$ . For some individuals this quantity will be positive, indicating that they are closer to group  $j$  than  $k$ , while for others it will be negative, if they are closer to  $k$  than  $j$ . Hence, the variation in the relative distance to group  $j$  and  $k$  can be quantified by the standard deviation of  $\Delta_{jk}(i)$  calculated over the distribution of individual positions. For future reference, I denote this quantity as  $D_{jk}$ .<sup>10</sup> Notice that  $D_{jk} = 0$  if and only if all individuals are equally distant to groups  $j$  and  $k$ . In substantive terms, this means that the rate at which group  $j$  appears in any of the individuals' networks will be equal to the rate at which group  $k$  appears, after adjustments for the gregariousness of individuals and the population shares of the groups.  $D_{jk}$  measures the deviation from this state: the higher the standard deviation, the more frequently will only one or the other group appear in individuals' acquaintanceship networks, but not both simultaneously. Therefore, I take  $D_{jk}$  as a direct measure of how much networks are segregated with respect to the groups  $j$  and  $k$ .

---

<sup>10</sup>Using the median absolute deviation (MAD), instead of the standard deviation, as an estimator of the variability in distances did not change the substantive conclusion of the analysis. Results of the analysis using MAD can be found in the supplementary materials.

## KNOWN LIMITATIONS OF THE APPROACH

The approach to study segregation patterns proposed here shares several limitations with other methods that rely on self-reported ego-networks. First, it is not certain to what extent the reported acquaintanceship ties, which are based on individuals' perceptions of their networks, reflect actual patterns of interaction. Previous research on acquaintanceship ties suggests that individuals are better in recalling ties to rare social groups in the population (Killworth et al., 2003; Zheng et al., 2006). If this is also true for individuals' local environments, we might expect that the results of the following analysis are underestimating the degree of segregation in acquaintanceship networks. This is because the better individuals are able to remember acquaintances which are less common in their surroundings, the more heterogeneous their network will appear.

Second, the visibility of attributes differs across groups. Some attributes—such as ideology or sexual orientation—have to be disclosed by the alters before the respondent can report them, while others are easier to infer without such disclosure—such as race or being incarcerated. Hence, for attributes that are not directly visible, the estimates will suffer from selective disclosure bias. To the extent that individuals disclose their identity selectively to others whom they expect to agree with (Gerber et al., 2012; Cowan and Baldassarri, 2018), this would lead to overestimating the degree of network segregation. For example, gay and lesbian people might conceal their sexual orientation in front of acquaintances that they expect to be unwelcoming. Similarly, conservatives and liberals may hide their political identity in front of acquaintances who don't share their views in order to avoid disagreement. These tendencies will bias the reported acquaintanceship networks towards more homogeneity and, thus, more segregation with respect to these groups.

Third, respondents might differ in their interpretation of group attributes. This problem is particularly threatening for groups defined by religiosity and ideology. For example, attending religious services “rarely” or being “strongly liberal” might mean different things to different respondents (Hare et al., 2015; Park, 2018). Hence, even if two respondents know the same alter, she might be liberal for one respondent while conservative for the other. Previous research

on political networks suggests that individuals tend to project their own political views onto others, leading to an over-report of political homogeneity in self-reported networks (Huckfeldt et al., 1995, 2004). While such a tendency would lead to overestimating segregation based on ideology, it is not clear how much the analyzed data are affected by these biases.

Fourth, the analysis relies on external information about the population shares in order to interpret distances in social space as deviations from random mixing scenarios. The model tries to model measurement error by using prior distributions, rather than point estimates. But it is still possible that inaccurate population share estimates will bias the results. We might speculate that the between-group distances will be more influenced by mismeasurement of population shares than the directions in which groups are positioned vis-a-vis each other. This is because the mismeasurement of the population shares will affect the distances of a group to all individuals in the same way, so that underestimation of the population share will draw the group towards the cloud of individual positions, while overestimation will drive the group away from the cloud.

Lastly, while the measured groups were chosen to represent major potential cleavages in the contemporary U.S. (DiPrete et al., 2011), they hardly encompass all the divisions that influence the distribution of acquaintanceships. For example, there are other important aspects of socio-economic inequalities that are not captured by the categories Unemployed and Second-home Owners. In particular, the data lacks groups defined in terms of education, whereas the importance of college degrees for social standing has increased substantially during the last decades (Fischer and Hout, 2006) and is likely to influence with whom individuals become acquainted with. Another major group distinction that is not included in the data is partisanship, which has become increasingly salient in influencing not only political beliefs but also emotional antagonism and perceived social distances (Mason, 2018; Iyengar et al., 2019; but see, Klar et al., 2018). Hence, we will not be able to reach any conclusions regarding these characteristics.

Most of these limitations are inherent in analyses of sampled ego-network data and stem from the fact that we are unable to observe the alters directly and have to rely on ego's reports. A conservative way to interpret the results to follow would be, therefore, to see them as reflecting the perceptions of the public, rather than actual interaction patterns (DiPrete et al., 2011). In

other words, the results should be understood as showing segregation patterns in “reported” networks. Despite these limitations, the model offers us a new way to look into sampled ego-centric network and gain insights into the segregation patterns of individuals’ broader network along multiple social dimensions simultaneously. Further, while this limitations of the data should be kept in mind, relying on sampled ego-networks is often the only way to study tie distributions in large collectivities, where obtaining data on the interaction pattern of the whole population is impossible.

## Results

### MODEL COMPARISON

As with other latent variable models, I first determine how many dimensions are needed to account for the pattern of relationships that link respondents to the thirteen groups. In order to get a sense of the predictive fit of a model, I rely on two measures: the WAIC and the PSIS-LOOIC (Vehtari et al., 2017). Both WAIC and LOOIC are approximations to the out-of-sample deviance and are asymptotically equivalent. Smaller values indicate a better predictive fit of the model. Starting with a 2-dimensional social space model, up to 5 dimensions were fitted to the data. Results of the analysis are presented in Table 2. For comparison, the predictive fit of the random mixing model and Negative Binomial model with group-specific dispersion parameters is shown as well.<sup>11</sup> The lower LOOIC and WAIC values suggest that the social space models outperform both the random mixing as well as the group-specific dispersion model. Further, both information criteria attain their lowest value for the three-dimensional social space model, suggesting that it performs the best. Examinations of the posterior predictive distribution, shown in the supplementary materials, suggest that the three-dimensional model performs fairly well in following the general trend in the data. Thus, the three-dimensional model will

---

<sup>11</sup>This model is similar but not identical to the model used in DiPrete et al. (2011). The major difference is that the  $\pi_j$  parameters are given informative priors outlined in Table 1, rather than being estimated from in-degrees of the groups. This will lead to a poorer fit of the model, since the fit is almost entirely determined by the gregariousness parameters,  $\gamma_i$ , and the dispersion parameters,  $\phi_j$ .

Table 2: Model Comparison Based on Leave-One-Out Cross Validation using Pareto-Smoothed Importance Sampling (LOOIC) and Watanabe-Akaike Information Criterion (WAIC)

Model	LOOIC	WAIC
Social Space, 5 Dimensions	19174.5	19006.2
Social Space, 4 Dimensions	19169.8	18969.2
Social Space, 3 Dimensions	19064.4	18828.1
Social Space, 2 Dimensions	19088.6	18874.1
Group-specific Dispersion <sup>a</sup>	19604.5	19581.5
Random Mixing	29796.7	29708.9

<sup>a</sup> Negative-Binomial model with group-specific dispersion parameter.

be used in the analysis that follows.<sup>12</sup>

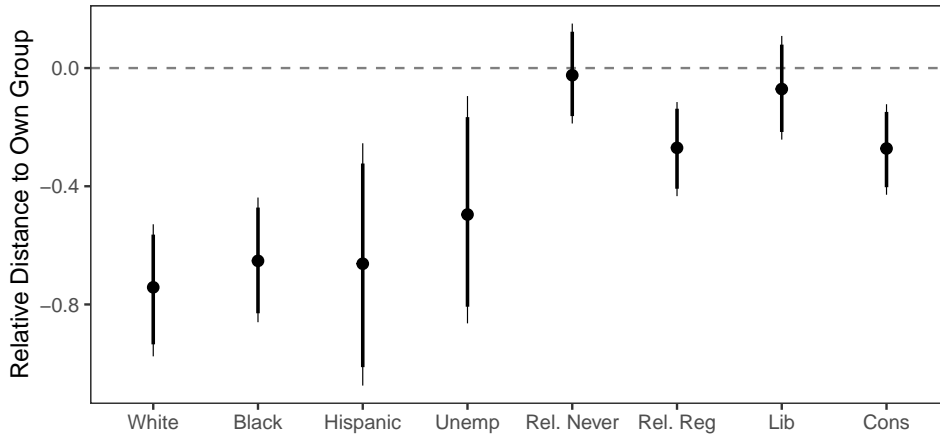
## SALIENCE OF GROUP ATTRIBUTES

Figure 3 shows the association between group membership and social distances in the estimated social space. The  $y$ -axis represents the difference in the median distance to a social group between respondents who belong to the group and those who do not. The black points represent the posterior medians and the thick and thin vertical lines, respectively, the 90% and 95% credible intervals. In general, the figure suggests that not only close ties of core networks (Marsden, 1987; Smith et al., 2014) but also reports of acquaintanceship ties tend to be homophilous. For all groups, the distance of individuals to their own groups tends to be smaller than that for individuals who do not belong to it, although for the Strongly Liberal (Lib) and the Non-religious group (Rel. Never) the estimates are close to zero.

The strongest in-group prevalences are found for the racial groups, which is in line with previous research on homophily (McPherson et al., 2001; Moody, 2001; Smith et al., 2014). For example, Whites, Blacks, and Hispanics are all estimated to be closer, on average, by 0.7 units in social space to their own group than out-group members. To put this number into perspective, for an individual who reports 50 acquaintanceship ties to members of a group when being three units apart from that group, reducing the distance by 0.7 would imply that she would report approximately 70% more ties, so that she'll end up with about 85. For individuals

<sup>12</sup>It should be noted that the model is better in predicting the responses indicating above 10 acquaintanceship ties than those indicating lower numbers. This seems to be due to the fact that many respondents have at least one acquaintance in most of the groups, so that the highest category gives the strongest signal about their social position. See Figure S1 in the online appendix for detailed results.

Figure 3: Saliency of Social Groups



Notes: 1) Y-axis represents the difference in the median social distance to groups between individuals belonging to the group and individuals who do not. 2) Thin and thick vertical bars show, respectively, the 95% and 90% posterior intervals of the differences; dots represent the posterior median. 2) Individuals who self-identified as “liberal” or “extremely liberal” are used as the in-group for the “Strongly liberal” category (Lib); the in-group for the “Strongly conservative” category (Cons) is defined analogously. For the “Attending religious services regularly” category (Rel. Reg), individuals who report to attend services at “nearly every week” or more are used as the in-group; for the “Attend rarely/never” category (Rel. Never), respondents who report to attend services “once a year” or less are used.

who are already close to the group, the predicted increase in the number of reported ties by moving them closer by 0.7 units would be even larger.<sup>13</sup> Thus, the model suggests that racial homophily in reported acquaintanceship networks is quite substantial. On the other hand, homophily based on ideology and religiosity seems to be relatively weak when compared to that based on race, once gregariousness and population distributions are taken into account.<sup>14</sup> Yet, homophily is an inherently binary concept in that it compares tie-distributions on single social dimensions in isolation. To examine how acquaintanceship ties are distributed across multi-

<sup>13</sup>The expected number of reported ties that connects individual  $i$  to group  $j$  under the model is  $\mu_{ij} = \gamma_i \pi_j \rho_{ij}^{-2}$ . Thus, the percentage increase in the number of ties to the group when the distance is reduced by the amount of  $\Delta$  is  $100 \times \left[ \frac{\gamma_i \pi_j (\rho_{ij} - \Delta)^{-2} - \gamma_i \pi_j \rho_{ij}^{-2}}{\gamma_i \pi_j \rho_{ij}^{-2}} \right] = 100 \times [(1 - \Delta/\rho_{ij})^{-2} - 1]$ . Plugging-in  $\Delta = .7$  and  $\rho_{ij} = 3$  gives approximately 70.13. As the percentage increase is a decreasing function of  $\rho_{ij}$ , it follows that the increase in ties to group  $j$  will be larger if  $\rho_{ij} < 3$  for the same reduction  $\Delta$ .

<sup>14</sup>It is also interesting to note that the in-group biases tend to be asymmetric. For example, respondents who are conservative and attend religious services regularly tend to be closer to their own groups, while such tendencies among liberals and respondents who attend religious services rarely or never is rather weak and estimated to be close to zero. A similar asymmetry in political homophily was found in online social networks (Barberá et al., 2015; Boutyline and Willer, 2017), suggesting that the stronger tendency of conservatives to associate with like-minded others is not a particularity of the analyzed data.



ple social dimensions simultaneously, we need to examine not only in-group distances but the structure of the whole social space. This is where we turn next.

## POSITIONS OF INDIVIDUALS AND GROUPS IN SOCIAL SPACE

The estimated joint space of individuals and groups is shown in Figure 4. Although the labeling of the dimensions is immaterial and only distances in the space are meaningful, the space is rotated such that the first two dimensions are aligned with the directions along which individuals' positions vary the most to help interpretation. Panel (a) of the figure shows the posterior median positions of groups and individuals in the estimated three-dimensional space, where groups are represented by black dots and respondents by gray dots. Panel (b) shows posterior medians of the between-group distances.<sup>15</sup>

By considering multiple groups simultaneously in the space, we see that the largest social distance is not found on single dimensions of differentiation—such as race, religiosity, or ideology—but in combinations across them. For example, the distance between Blacks and Second-home Owners (post. median: 6.68, 95% cred. int: [5.52, 7.96]) exceeds the distance between Blacks and Whites (post. median: 3.83, 95% cred. int: [3.06, 4.78]) or Regular Churchgoers and the Non-religious group (post. median: 3.84, 95% cred. int: [2.86, 4.82]). These results suggest that the deepest cleavages in terms of co-appearance rates of groups in reported acquaintanceship networks lies not along the racial dimension nor the economic dimension, but across them. As discussed above, these estimates will be sensitive to the estimated population proportions and require caution. Still, these results suggest that studying social distance on each social dimension in isolation, while being important in itself, might fail to capture other divisions that separate the population more than any single dimension alone.

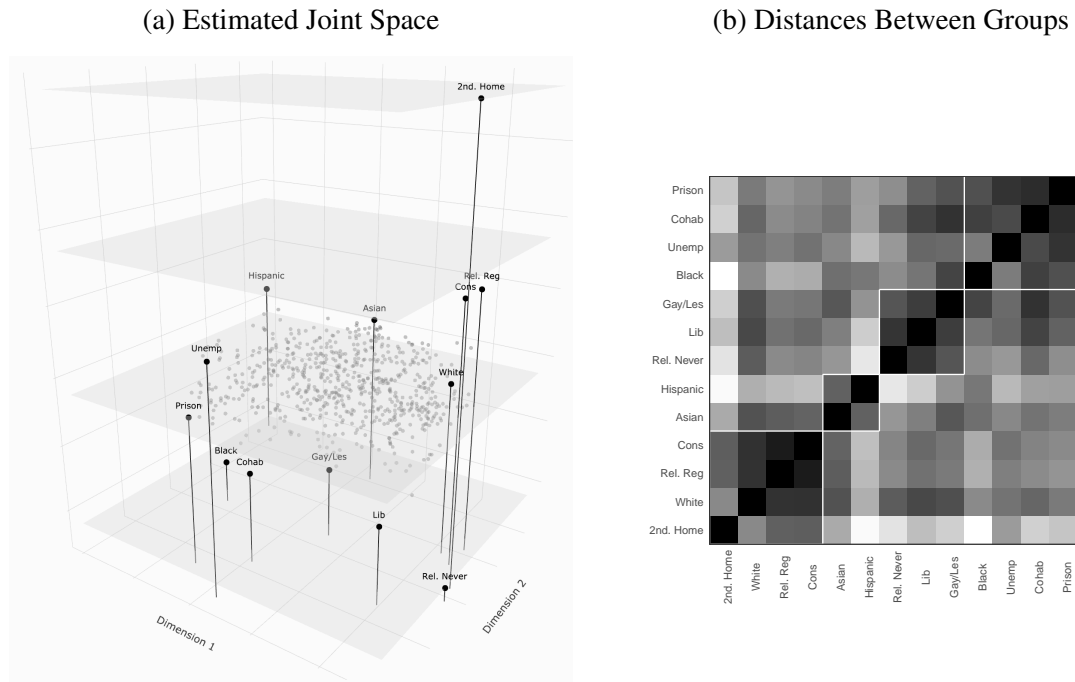
When it comes to the positions of individuals, it turns out that most of them are distributed along a plane parallel to the first two dimensions, while showing less variation across the third dimension.<sup>16</sup> Since segregation patterns, as the term is defined here, crucially depend on the variation of individual positions, the first two dimensions of the space are shown separately in

---

<sup>15</sup>The thick white lines represents clustering results from applying Gaussian mixture models to the posterior medians and choosing the model with the lowest Bayesian Information Criterion.

<sup>16</sup>About 92% of the variance in the positions of the respondents is accounted for by the first two dimensions.

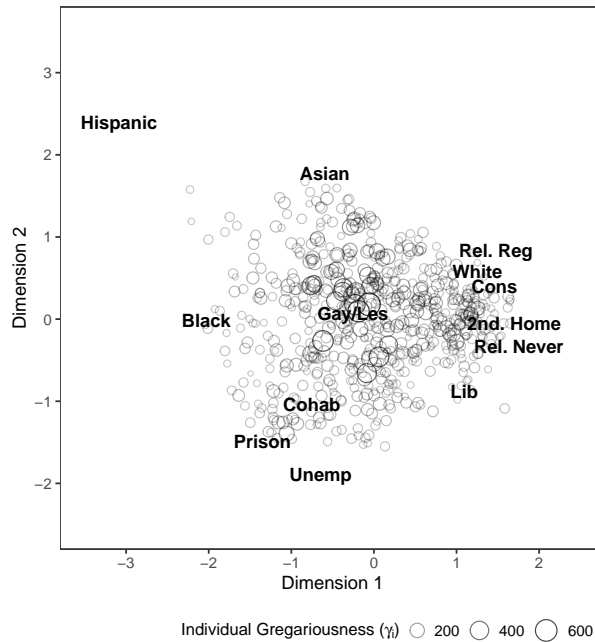
Figure 4: Positions of Respondents and Groups in Social Space



Notes: 1) Posterior medians of group positions are shown as black dots in the plot on the left; posterior medians of individual positions are represented by grey dots. 2) Plot on the right shows the posterior median of the between-group distances. Thick white lines represent the clustering results from applying a Gaussian mixture model to the posterior medians of the group-positions.

Figure 5. Recall that the space is rotated such that the first two dimensions are the directions along which individual positions, and hence the group-composition of their networks, vary the most. The figure shows these two dimensions reflect, although not perfectly, socioeconomic inequalities across the groups, with relatively disadvantaged groups being placed closer to the lower-left corner of Figure 5. The third dimension, on the other hand, differentiates the position of groups that are placed on the right-side of Dimension 1 (see Figure 4). The lack of variation in individuals' position along the third dimension suggest that individuals are either close these group or far from them, rather than some individuals being close to one and distant to the other. For example, while Strongly Liberal and Conservative group are not necessarily close in social space, most individuals who are close to Strongly Liberal group tend to be close to the Strongly Conservative group as well and, hence, network are not strongly segregated across these groups. This is not the case for the pair of groups {Black, White} or {Unemployed,

Figure 5: Distribution of Respondents in Social Space, First Two Dimensions



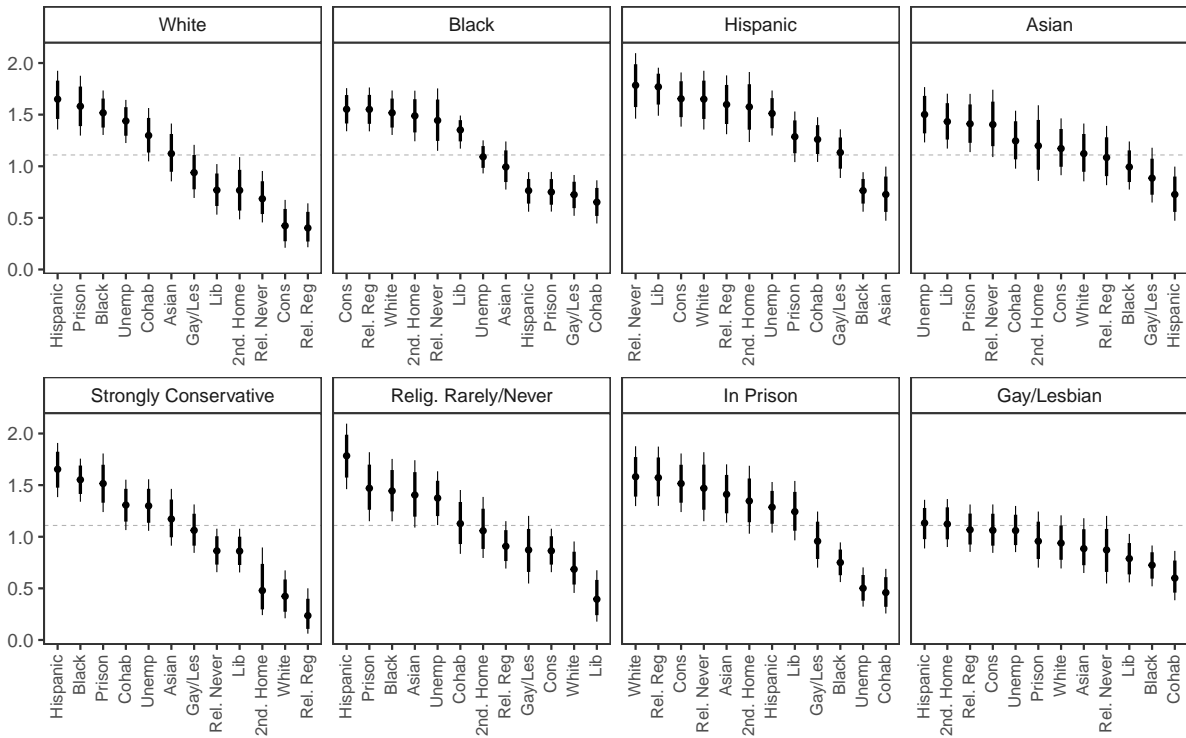
Notes: a) The first two dimensions account for 91.5% of the variance in the respondents' posterior median positions. The corresponding number of the group-position is 82.4%. b) The sizes of the circles are proportional to the posterior median gregariousness ( $\gamma_i$ ) of the respondents.

Asian}, where individuals close to one group tend to be far from the other, suggesting that networks containing many alters of one group tend to contain few of the other group.

To examine these claims more formally,  $D$ -values were calculated and compared across pairs of groups. Posterior medians, together with the 90% and 95% credible intervals, of these estimates are shown in Figure 6 for a subset of the groups.<sup>17</sup> These values are calculated over all three dimensions of the social space. The results confirm the intuition that between-group distances are not sufficient to describe the lines along which reported acquaintanceship networks are segregated. For example, while the Non-religious group and the Second-home Owner group occupy opposite poles on the third dimension of the space, their estimated  $D$ -value is 1.06 (95% cred. int: [0.79, 1.39]), which is close to the average value across all pairs of groups ( $\bar{D} = 1.11$ ). On the other hand, Blacks and Whites, for which the distance is smaller than that of Non-religious and Second-home Owners groups, generate larger variations in network compositions with  $D = 1.52$  (95% cred. int = [1.30, 1.73]). The smallest value of

<sup>17</sup>The full set of results are available in Figure S3 of the supplementary materials.

Figure 6: Variation in Individuals' Distances to Pairs of Groups



Notes: Dots represent posterior medians of estimated  $D$  values. Thin and thick vertical lines, respectively, show the 90% and 95% credible intervals. Dashed horizontal line in gray shows the posterior median of the average  $D$  values across all pairs of groups.

$D$  is found between Strongly Conservatives and Regular Churchgoers ( $D = .24$ , 95% cred. int =  $[0.06, 0.50]$ ), implying that individuals tend to be either simultaneously close or distant from these groups, but not close to one group and distant to the other.

The  $D$ -values illustrate as well that the first two dimensions of the space are most important to understand the segregation patterns of reported acquaintanceship networks. Figure 6 shows that segregation is relatively modest between pairs of groups that are both placed on the left- or on the right-side of Dimension 1, when compared to segregation between groups that lie on opposite sides of the dimension. For example, Whites shows strong segregating tendencies when paired with Hispanics, Prisoners, Blacks, Unemployed, and Cohabiting Women, while relatively little when paired with Strong Liberals, Strong Conservatives, Second-home Owners, Non-religious, and Regular Churchgoers. A similar pattern is observed for the Strongly Conservative group or the Non-religious group. For Blacks, on the other hand, we see almost

the exact opposite pattern. Thus, Figure 6 shows that the main cleavages apparent in Figure 5, which showed only the first two dimensions, are also found when variations in individual and group positions over all three dimensions are taken into account. Indeed, the pairs {Strongly Conservative, Strongly Liberal} and {Regular Churchgoers, Non-religious} with posterior median  $D$  values equal to 0.86 and 0.91, respectively, are segregating networks less than the racial groups for which  $D$  is estimated to be 1.12 for {Whites, Asians}, 1.52 for {Whites, Blacks}, and 1.65 for {Whites, Hispanics}, even though the between-group distance between Regular Churchgoers and Non-religious is on par with that of Blacks and Whites.

One group that stands out among all is the Gay and Lesbian group which is positioned close to the center of the social space and for which  $D$  is estimated to be either close to or lower than the average  $D$  across all groups. To over simplify a bit, this shows the tendency of gay and lesbian people to appear in reported acquaintanceship networks regardless of their composition in terms of race, socioeconomic status, ideology, or religiosity. Of course, there is some variability in that the  $D$  value is higher for the pairing with Second-home Owners than for the pairing with Cohabiting Women. In addition, gay and lesbian people might not be found in all networks, since network sizes tend to vary across individuals and the population share of the group is rather small. Still, among the groups under analysis, it appears that the Gay and Lesbian group comes closest to being randomly distributed across reported acquaintanceship networks. The implications of these findings will be discussed in the next section.

## Summary and Discussion

This paper introduced a model through which network segregation along multiple social dimensions can be analyzed simultaneously. As numerous theorists have argued, not only the depth of any social cleavage but the alignment of them defines how well society is integrated (Dahl, [1956]2006: Ch. 4; Coser, 1956: Ch. 4.; Lipset, 1960: Ch. 3; Blau, 1977). Similarly, more important than distances between any pairs of groups, or along any single dimension of social differentiation, is the constellation of the whole social space to understanding how network ties are binding, or fail to bind, society together. Most of the new insights gained from the

analysis would have been lost if the data were treated as consisting of isolated dyadic relationships, since they emerge only as the distances between all groups and individuals are examined simultaneously as interrelated to each other in a structured way.

The analysis of the 2006 GSS showed not only that reported acquaintanceship ties are far from randomly distributed in the U.S. but also that consolidated inequalities across multiple social dimensions are reflected in the distribution of these ties. Of course, segregation of the reported networks is far from perfect as individual positions vary continuously over whole the space rather than showing a deep unsurmountable split. This variation is important to acknowledge, since it shows that reported acquaintanceship networks are not strictly confined within specific groups and show considerable amounts of heterogeneity. Still, that relatively disadvantaged and well-off groups were found to occupy distinct areas of the space suggests that inequality consolidated across several social dimensions is shaping the distribution of not only strong but also weaker ties, which are often hoped to be less influenced by such structural tendencies. It should be kept in mind that the results are based on reported networks which might differ from the actual acquaintanceship patterns of the U.S. population. Yet, even if the captured structure corresponds to pure perceptions, these perceptions will have real-world consequences and are, therefore, important to understand.

In contrast to previous research that found that acquaintanceship ties are segregated across virtually all social dimensions (DiPrete et al., 2011), the analysis presented here suggests that race is the most predominant cleavage across which these ties have difficulties to span. Not only are homophilous tendencies the strongest for the racial groups, but they tend to segregate reported acquaintanceship networks to a higher degree than ideology or religiosity. Indeed, while groups defined in terms of ideology and, especially, religiosity are not necessarily close in social space, individuals tend to be either close or distant from both groups, rather than being close to one and distant from the other.<sup>18</sup>

---

<sup>18</sup>Black and White respondents are found to be concentrated around different centers in the estimated space as well. This suggests that Blacks and Whites not only receive acquaintanceship ties from different sources as a group but that their ego-network compositions themselves tend to differ, with Blacks having relatively more ties to the positioned on the left of the first dimension and Whites more ties to the groups on the opposite side of the space. Distinctions based on family income, ideology, and religiosity, on the other hand, do not differentiate the distributions of respondents in the space as well. Details on the additional analyses can be found in the supplementary materials.

The relatively modest amount of segregation found between the ideological groups might surprise some readers, as the emotional antagonism between liberals and conservatives is well-documented and was already at a high level in 2006 (Iyengar et al., 2012; Mason, 2018; Iyengar et al., 2019). Yet, as studies have repeatedly shown, even political discussion networks which usually consist of relatively “strong” ties (Klofstad et al., 2009) tend to be not as segregated by political ideology as it is often thought (Huckfeldt et al., 2004); and difference in ideology might not be a major concern when individuals decide whether to create or dissolve acquaintanceship ties that are relatively weak. In addition, we often cannot know the political orientation of our acquaintances before we become acquainted to them. Severing a social tie altogether after it has been created due difference in political views would be quite an extreme decision for many individuals, given that they can always “keep their distance” while remaining an acquaintance. This is especially true for relationships to alters with whom individuals have to interact on a regular basis in their everyday lives.<sup>19</sup> Lastly, that the distance between the ideological groups is smaller than their distance to groups occupying the opposite side of the social space suggests that network compositions of reported acquaintanceship ties are not as clearly split between liberal- or conservative-dominated networks as they are between networks that contain ideologues of both convictions and networks containing no ideologues at all. Hence, the distribution of these ties appears to reflect political inequality rather than political polarization: namely, the division between individuals embedded in resource-rich political environments—where people are politically engaged, political identities are salient, and politics is routinely discussed—and those embedded in contexts where political resources are scarce (Verba et al., 1995; Schlozman et al., 2012).

Another finding that emerged from the analysis is that the Gay and Lesbian group was placed near the center of the estimated space. This is remarkable since gay rights issues lied at the core of partisan conflict during the time the data were collected: the difference in support for gay rights between non-black Democrats and Republicans in the mid 2000s was rivaling the difference between northerners and southerners on civil rights issues in the 1960s, and twenty-

---

<sup>19</sup>These results are also in line with findings on online social networks, where it has been shown that non-political events, such as the Super Bowl or Winter Olympics, were “retweeted” across the ideological spectrum (Barberá et al., 2015), implying that many liberals and conservatives are already connected to (i.e., “following”) each other.

one states held referenda to ban same-sex marriage between 2004 and 2006 (Hetherington, 2009). Yet, at the same time as gay rights issues became increasingly divisive, gay and lesbian people were one of the best integrated groups in reported acquaintanceship networks in the U.S. Of course, being positioned near the center of the social space does not mean that this group received the most incoming ties. Rather, it implies that individuals with very different network compositions were similarly likely to report that they know someone who is gay. We might speculate that it was this unique position, ideal for diffusion of new ideas, that provided the structural backbone behind the unrivaled rapidity at which support for gay rights surged over the last thirty years (Baldassarri and Park, 2020).<sup>20</sup>

In sum, this paper offers researchers new conceptual and analytical tools to study network segregation. Yet, the utility of the tools developed here is not confined to the analysis of networks. The very concept of a “space” attunes us to the relative nature of various social phenomena, such as segregation, inequality, or polarization. In particular, it pushes us to escape the often implicit view that societies consist of sharply demarcated boundaries without an explicit formulation of how the clusters implied by them are positioned vis-a-vis one another. To be sure, social boundaries exist; but probably more often in our mind than outside of it (Rosch and Lloyd, 1978; Freeman and Webster, 1994; Zerubavel, 1997). Even the sharpest boundaries are permeable, which leads to a continuous social landscape with regions of higher and lower density rather than clearly segregated clusters or categories. A spatial representation enables the description of such subtleties as well as the detection of social boundaries if they are present. In this regard, both the conceptual and methodological tools developed here could be applied to other topics, where the relative, rather than absolute, position of individuals, groups, or other objects are of theoretical importance.

## References

Abramowitz, Alan. 2010. *The Disappearing Center*. CT: Yale University Press.

Baldassarri, Delia and Andrew Gelman. 2008. “Partisans without constraint: Political polarization and trends in American public opinion.” *American Journal of Sociology* 114:408–446.

---

<sup>20</sup>Indeed, a recent analysis of the same dataset used here found that being acquainted with gay or lesbian people was associated with increased support for gay rights issues four years later (DellaPosta, 2018).



- Baldassarri, Delia and Barum Park. 2020. "Was There a Culture War? Polarization and Secular Trends in US Public Opinion." *The Journal of Politics* 82:809–827.
- Barberá, Pablo, John T Jost, Jonathan Nagler, Joshua A Tucker, and Richard Bonneau. 2015. "Tweeting from left to right: Is online political communication more than an echo chamber?" *Psychological Science* 26:1531–1542.
- Bearman, Peter and Paolo Parigi. 2004. "Cloning headless frogs and other important matters: Conversation topics and network structure." *Social Forces* 83:535–557.
- Bishop, Bill. 2009. *The big sort*. NY: Mariner Books.
- Blau, Peter M. 1977. *Inequality and Heterogeneity*. NY: Free Press.
- Blau, Peter M. 1994. *Structural contexts of opportunities*. University of Chicago Press.
- Blau, Peter M and Joseph E Schwartz. [1984]1997. *Crosscutting Social Circles*. NJ: Transaction.
- Borg, Ingwer and Patrick JF Groenen. 2005. *Modern multidimensional scaling*. NY: Springer.
- Borgatti, Stephen P and Martin G Everett. 1997. "Network analysis of 2-mode data." *Social networks* 19:243–269.
- Bottero, Wendy and Kenneth Prandy. 2003. "Social interaction distance and stratification." *British Journal of Sociology* 54:177–197.
- Bourdieu, Pierre. 1989. "Social space and symbolic power." *Sociological Theory* 7:14–25.
- Boutyline, Andrei and Robb Willer. 2017. "The social structure of political echo chambers: Variation in ideological homophily in online networks." *Political Psychology* 38:551–569.
- Brashears, Matthew E, Michael Genkin, and Chan S Suh. 2017. "In the organization's shadow: How individual behavior is shaped by organizational leakage." *American Journal of Sociology* 123:787–849.
- Breiger, Ronald L. 1974. "The duality of persons and groups." *Social Forces* 53:181–190.
- Burt, Ronald S. 1984. "Network items and the general social survey." *Social Networks* 6:293–339.
- Burt, Ronald S. 1990. "Kinds of Relations in American Discussion Networks." pp. 411–453, In *Structures of Power and Constraint*, edited by Mayer Marshall W. Scott W. Richard Calhoun, Craig. NY: Cambridge University Press.
- Centola, Damon. 2015. "The Social Origins of Networks and Diffusion." *American Journal of Sociology* 120:1295–1338.
- Coser, Lewis A. 1956. *The Functions of Social Conflict*. NY: Free Press.
- Cowan, Sarah K and Delia Baldassarri. 2018. "'It could turn ugly': Selective disclosure of attitudes in political discussion networks." *Social Networks* 52:1–17.
- Dahl, Robert A. [1956]2006. *A Preface to Democratic Theory*. IL: University of Chicago Press.

- Davis, James A. 1967. "Clustering and Structural Balance in Graphs." *Human Relations* 20:181–187.
- Davis, James A. 1970. "Clustering and hierarchy in interpersonal relations: Testing two graph theoretical models on 742 sociomatrices." *American Sociological Review* pp. 843–851.
- DellaPosta, Daniel. 2018. "Gay Acquaintanceship and Attitudes toward Homosexuality: A Conservative Test." *Socius* 4:1–12.
- DellaPosta, Daniel, Yongren Shi, and Michael Macy. 2015. "Why do liberals drink lattes?" *American Journal of Sociology* 120:1473–1511.
- DiMaggio, Paul, John H. Evans, and Bethany Bryson. 1996. "Have American's social attitudes become more polarized?" *American Journal of Sociology* 102:690–755.
- DiPrete, Thomas A, Andrew Gelman, Tyler McCormick, Julien Teitler, and Tian Zheng. 2011. "Segregation in social networks based on acquaintanceship and trust." *American Journal of Sociology* 116:1234–83.
- Fararo, Thomas J. 1992. *The meaning of general theoretical sociology*. NY: Cambridge University Press.
- Fiorina, Morris P. and Samuel J. Abrams. 2008. "Political Polarization in the American Public." *Annual Review of Political Science* 11:563–588.
- Fischer, Claude S and Michael Hout. 2006. *Century of difference*. NY: Russell Sage Foundation.
- Fischer, Claude S. and Greggor Mattson. 2009. "Is America Fragmenting?" *Annual Review of Sociology* 35:435–455.
- Freeman, Linton C. 1978. "Segregation in social networks." *Sociological Methods & Research* 6:411–429.
- Freeman, Linton C. and Cynthia M. Webster. 1994. "Interpersonal Proximity in Social and Cognitive Space." *Social Cognition* 12:223–247.
- Gates, Gary J. 2011. *How many people are lesbian, gay, bisexual, and transgender?* CA: Williams Institute.
- Gerber, Alan S, Gregory A Huber, David Doherty, and Conor M Dowling. 2012. "Disagreement and the avoidance of political discussion: Aggregate relationships and differences across personality traits." *American Journal of Political Science* 56:849–874.
- Granovetter, Mark S. 1973. "The strength of weak ties." *American Journal of Sociology* 78:1360–1380.
- Greenacre, Michael and Trevor Hastie. 1987. "The geometric interpretation of correspondence analysis." *Journal of the American statistical association* 82:437–447.
- Hare, Christopher, David A. Armstrong, Ryan Bakker, Royce Carroll, and Keith T. Poole. 2015. "Using Bayesian Aldrich-McKelvey Scaling to Study Citizens' Ideological Preferences and Perceptions." *American Journal of Political Science* 59:759–774.

- Hetherington, Marc J. 2009. "Putting Polarization in Perspective." *British Journal of Political Science* 39:413.
- Hoff, Peter D, Adrian E Raftery, and Mark S Handcock. 2002. "Latent space approaches to social network analysis." *Journal of the American Statistical Association* 97:1090–1098.
- Hofstra, Bas, Rense Corten, Frank Van Tubergen, and Nicole B Ellison. 2017. "Sources of segregation in social networks: A novel approach using Facebook." *American Sociological Review* 82:625–656.
- Hout, Michael and Claude S Fischer. 2014. "Explaining Why More Americans Have No Religious Preference: Political Backlash and Generational Succession, 1987-2012." *Sociological Science* 1:423–447.
- Huber, Gregory A and Neil Malhotra. 2017. "Political homophily in social relationships: Evidence from online dating behavior." *The Journal of Politics* 79:269–283.
- Huckfeldt, Robert, Paul Allen Beck, Russell J. Dalton, and Jeffrey Levine. 1995. "Political Environments, Cohesive Social Groups, and the Communication of Public Opinion." *American Journal of Political Science* 39:1025.
- Huckfeldt, Robert, Paul E Johnson, and John Sprague. 2004. *Political disagreement*. NY: Cambridge University Press.
- Hunter, James Davison. 1991. *Culture Wars*. NY: Basic Books.
- Iyengar, Shanto, Yphtach Lelkes, Matthew Levendusky, Neil Malhotra, and Sean J Westwood. 2019. "The origins and consequences of affective polarization in the United States." *Annual Review of Political Science* in press.
- Iyengar, S., G. Sood, and Y. Lelkes. 2012. "Affect, Not Ideology: A Social Identity Perspective on Polarization." *Public Opinion Quarterly* 76:405–431.
- Iyengar, Shanto and Sean J. Westwood. 2015. "Fear and Loathing across Party Lines: New Evidence on Group Polarization." *American Journal of Political Science* 59:690–707.
- Kaufmann, Karen M. 2002. "Culture wars, secular realignment, and the gender gap in party identification." *Political Behavior* 24:283–307.
- Killworth, Peter D, Christopher McCarty, H Russell Bernard, Eugene C Johnsen, John Domini, and Gene A Shelley. 2003. "Two interpretations of reports of knowledge of subpopulation sizes." *Social Networks* 25:141–160.
- Klar, Samara, Yanna Krupnikov, and John Barry Ryan. 2018. "Affective polarization or partisan disdain? Untangling a dislike for the opposing party from a dislike of partisanship." *Public Opinion Quarterly* 82:379–390.
- Klofstad, Casey A, Scott D McClurg, and Meredith Rolfe. 2009. "Measurement of political discussion networks: A comparison of two "name generator" procedures." *Public Opinion Quarterly* 73:462–483.
- Klofstad, Casey A, Rose McDermott, and Peter K Hatemi. 2013. "The dating preferences of liberals and conservatives." *Political Behavior* 35:519–538.

- Laumann, Edward O and Louis Guttman. 1966. "The relative associational contiguity of occupations in an urban setting." *American Sociological Review* pp. 169–178.
- Laumann, Edward O. and Franz U. Pappi. 1976. *Networks of collective action*. NY: Academic Press.
- Lipset, Seymour M. 1960. *Political Man*. NY: Doubleday.
- Lorrain, Francois and Harrison C White. 1971. "Structural equivalence of individuals in social networks." *The Journal of Mathematical Sociology* 1:49–80.
- Marsden, Peter V. 1987. "Core discussion networks of Americans." *American sociological review* pp. 122–131.
- Marsden, Peter V. 1988. "Homogeneity in confiding relations." *Social Networks* 10:57–76.
- Martin, John Levi. 2009. *Social structures*. NJ: Princeton University Press.
- Mason, Lilliana. 2018. *Uncivil agreement*. IL: University of Chicago Press.
- McCarty, Nolan, Keith T Poole, and Howard Rosenthal. 2006. *Polarized America*. MA: MIT Press.
- McCormick, Tyler H and Tian Zheng. 2015. "Latent surface models for networks using Aggregated Relational Data." *Journal of the American Statistical Association* 110:1684–1695.
- McPherson, J. Miller and James R. Ranger-Moore. 1991. "Evolution on a dancing landscape: organizations and networks in dynamic Blau space." *Social Forces* 70:19–42.
- McPherson, Miller. 2004. "A Blau space primer: prolegomenon to an ecology of affiliation." *Industrial and Corporate Change* 13:263–280.
- McPherson, Miller, Lynn Smith-Lovin, and Matthew E Brashears. 2006. "Social isolation in America: Changes in core discussion networks over two decades." *American Sociological Review* 71:353–375.
- McPherson, Miller, Lynn Smith-Lovin, and James M Cook. 2001. "Birds of a feather: Homophily in social networks." *Annual Review of Sociology* pp. 415–444.
- Moody, James. 2001. "Race, school integration, and friendship segregation in America." *American Journal of Sociology* 107:679–716.
- Mummolo, Jonathan and Clayton Nall. 2017. "Why partisans do not sort: The constraints on political segregation." *The Journal of Politics* 79:45–59.
- Newman, Mark EJ and Juyong Park. 2003. "Why social networks are different from other types of networks." *Physical Review E* 68:036122.
- Park, Barum. 2018. "How Are We Apart? Continuity and Change in the Structure of Ideological Disagreement in the American Public, 1980–2012." *Social Forces* 96:1757–1784.
- Park, Robert E. 1915. "The City: Suggestions for the Investigation of Human Behavior in the City Environment." *American Journal of Sociology* 20:577–612.

- Reardon, Sean F and Kendra Bischoff. 2011. "Income inequality and income segregation." *American Journal of Sociology* 116:1092–1153.
- Rivera, Mark T, Sara B Soderstrom, and Brian Uzzi. 2010. "Dynamics of dyads in social networks: Assortative, relational, and proximity mechanisms." *Annual Review of Sociology* 36:91–115.
- Rosch, Eleanor and Barbara Bloom Lloyd (eds.). 1978. *Cognition and categorization*. NJ: Lawrence Erlbaum Associates.
- Schlozman, Kay Lehman, Sidney Verba, and Henry E Brady. 2012. *The unheavenly chorus*. NJ: Princeton University Press.
- Schwartz, Christine R. 2013. "Trends and variation in assortative mating: Causes and consequences." *Annual Review of Sociology* 39:451–470.
- Simmel, Georg. 1908. *Soziologie*. Leipzig: Duncker & Humblot.
- Skvoretz, John. 1983. "Salience, Heterogeneity and Consolidation of Parameters: Civilizing Blau's Primitive Theory." *American Sociological Review* 48:360.
- Smith, Jeffrey A. 2017. "A Social Space Approach to Testing Complex Hypotheses: The Case of Hispanic Marriage Patterns in the United States." *Socius* 3:2378023117739176.
- Smith, Jeffrey A., Miller McPherson, and Lynn Smith-Lovin. 2014. "Social distance in the United States: Sex, race, religion, age, and education homophily among confidants, 1985 to 2004." *American Sociological Review* 79:432–456.
- Sorokin, Pitirim A. [1927]1959. *Social Mobility*. Reprinted in Pitirim A. Sorokin. 1959. *Social and Cultural Mobility*. NY: Free Press.
- Vehtari, Aki, Andrew Gelman, and Jonah Gabry. 2017. "Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC." *Statistics and Computing* 27:1413–1432.
- Verba, Sidney, Kay Lehman Schlozman, and Henry E Brady. 1995. *Voice and equality*. MA: Harvard University Press.
- Wasserman, Stanley and Katherine Faust. 1994. *Social network analysis*. NY: Cambridge university press.
- Watts, Duncan J., Peter Sheridan Dodds, and Mark EJ Newman. 2002. "Identity and search in social networks." *Science* 296:1302–1305.
- Wimmer, Andreas and Kevin Lewis. 2010. "Beyond and below racial homophily: ERG models of a friendship network documented on Facebook." *American Journal of Sociology* 116:583–642.
- Winship, Christopher. 1977. "A distance model for sociometric structure." *Journal of Mathematical Sociology* 5:21–39.
- Zerubavel, Eviatar. 1997. *Social mindscapes*. MA: Harvard University Press.
- Zheng, Tian, Matthew J Salganik, and Andrew Gelman. 2006. "How Many People Do You Know in Prison?: Using Overdispersion in Count Data to Estimate Social Structure in Networks." *Journal of the American Statistical Association* 101:409–423.