The Persistence of Division: Geography, Institutions, and Online Friendship Ties

Emma S. Spiro[1], Zack W. Almquist[2], and Carter T. Butts[3]

Abstract
As noted by theorists such as Blau, Durkheim, Mayhew, and others, interaction opportunity is a fundamental determinant of social structure. One of the most empirically well established factors influencing interaction opportunity is that of physical distance. The strength of this effect in modern societies, however, has been called into question because of technological advances (the so-called death-of-distance hypothesis). Here, the authors examine the effect of distance in an extreme case, considering weak friendship ties among university-affiliated persons in a large-scale online social network. Additionally, the authors explore institutional covariates, such as prestige and public or private status, as moderators of the relationship between distance and social tie probability. The findings demonstrate that geographical distance continues to affect social ties, despite the absence of physical barriers to tie formation and maintenance. The authors find moreover that institutional factors differentially affect the propensity for two university-affiliated individuals to be tied across large distances, illustrating that systematic differences in network structure along status lines persist even in ostensibly unconstrained settings.

Keywords
social networks, opportunity structures, higher education, spatial Bernoulli models, inequality

A basic insight of sociological theory is that *opportunity*—the possibility of interaction to occur—is a fundamental determinant of realized social structure (Blau 1977; Durkheim 1893; Mayhew 1984). Among others, factors such as physical distance, status, wealth, and institutional affiliation are known constraints on social interaction (Bossard 1932; Latané et al. 1995; Stewart 1941a). In the age of the Internet, however, the interplay between these traditional opportunity structures and social interaction has arguably changed; indeed, the current wave of technological change is only the latest in a long line of innovations that have altered the nature of human society (see, e.g., Benniger 1986; Bijker and Law 1994; Castells 1996; Diamond 1999; Pool 1977; Rheingold 2000; White 1964). New modes of transportation allow increased mobility, permitting social ties to be formed and maintained over greater distances. Advances in communication technologies change the time scale on which interpersonal communication over large distances is possible. More recently, technologies such as mobile phones and Internet-based messaging have made rapid communication across the globe commonplace. With the development and diffusion of these technologies comes the possibility that social interaction may no longer be restricted or constrained in the same ways that it had been

previously. Do traditional opportunity structures no longer shape social interaction? Or do these factors still persist, even in the face of radical technological change? This is the central question with which we are concerned.

Historically, one of the most important structural constraints on interaction has been physical distance (Bossard 1932; Caplow and Forman 1950; Festinger, Schachter, and Back 1950; Latané et al. 1995; Stewart 1941b; Zipf 1949). In recent years, however, many have questioned the continuing importance of proximity as a determinant of social interaction. This tradition is exemplified in the work of Cairncross (2001), who went so far as to promote the "death of distance" in modern society, claiming that communication technologies are rapidly obliterating distance as a relevant factor in

[1]University of Washington, Seattle, WA, USA
[2]University of Minnesota, Minneapolis, MN, USA
[3]University of California, Irvine, Irvine, CA, USA

Corresponding Author:
Emma S. Spiro, Information School, Department of Sociology,
CSSS, University of Washington, Mary Gates Hall Ste. 370, Seattle,
WA 98195, USA.
Email: espiro@uw.edu

determining not only how we live our day-to-day lives but how business is conducted around the globe.

Individuals in a social environment in which distance no longer affects social interaction are free to form social ties with myriad others across the globe. More important, the claim is made that if given this opportunity, they *will* do so. It is this assumption that underlies "death of distance" arguments. Although Cairncross's (2001) work is a popular representation of this hypothesis (see, relatedly, Friedman 2005), others have similarly incorporated death-of-distance themes into their own research, both explicitly and implicitly (Adams 1997; Kolko 2000; Wang, Lai, and Sui 2003). Indeed, numerous books with the title or theme of the "death of distance" have appeared in recent years (e.g., Cairncross 2001; O'Brien 1992; Vogelsang and Compaine 2000), receiving substantial public attention. Research in a variety of disciples within the social sciences contains similar themes; see, for example, the work of Glaeser and Ponzetto (2010) in economics, in which the authors argue that the "death of distance" is in part responsible for the decline of Detroit; the work of Castells (1996) in the field of communication, in which the author argues that the "death of distance" is changing the landscape of human communication and freeing the global labor force from geographical constraints; and the work of Wellman, Boase, and Chen (2002) and Rainie and Wellman (2012) in the field of sociology, in which the authors argue that space and distance are no longer limitations to creating a community. Mok, Wellman, and Carrasco (2010) even pose the titular question "Does Distance Still Matter in the Age of the Internet?" concluding that "e-mail contact is generally insensitive to distance."

Despite the considerable attention gleaned by "death of distance" claims, research in this area is not without dissenting voices. Some scholars disagree that communication and information technologies are in fact obliterating distance as a determinant of social interaction. Others believe the claims are exaggerated or premature. Onnela et al. (2011), for example, used data representing mobile phone calls and text messages to demonstrate that interaction probability for mobile phone contacts declines with distance.[1] Mesch, Talmud, and Quan-Haase (2012) considered self-reported instant message communication and found a similar distance effect on interaction. These results arguably suggest that technologies allowing contact at any distance have a limited effect on propinquity when they do not also provide users with a way to "discover" new social contacts at long distances. In social media environments, Takhteyev, Gruzd, and Wellman (2012) studied the impact of geography on Twitter relationships, concluding that many social ties lies within the same metropolitan region and that formal boundaries (e.g., national boarders) also predict tie probability. Likewise, Kulshrestha et al. (2012) showed that geography continues to have a significant impact on user interactions in the Twitter social network, suggesting that these effects could be explain by shared national, linguistic, and cultural backgrounds.

Backstrom, Sun, and Marlow (2010) also showed an empirical pattern of tie probability declining with physical distance, in this case for Facebook relationships. This prior work, however, tends to show an empirical relationship between tie probability and physical distance without modeling this association. Moreover, very few prior studies have examined factors moderating this relationship.

One of the challenges in evaluating the "death of distance" hypothesis lies in its preconditions. The theory maintains that in social environments in which individuals have access to unlimited numbers of potential social contacts, ties will be formed and exist without regard to geographic constraints. Inherent in this claim is the precondition that the social environment must afford individuals the opportunity to form and maintain social ties with whomever, wherever. In such a case, the observed marginal relationship between physical distance and social interaction is predicted to be flat. Although many Internet-based communication and information technologies facilitate interaction across large distances, not all tools actually afford users the opportunity and ability to both initiate and maintain social ties with the global population at large; an effective test of the hypothesis must thus focus on a setting that provides these capabilities.

Online social networking sites (OSNs) are a key Internet-based communication medium that frequently *does* allow users to browse public or semipublic profiles of other users, "network" (i.e., initiation relationships with known contacts and strangers), and articulate their social ties (Ellison 2007) irrespective of geographical location. OSNs expand the population of potential contacts, allowing the opportunity to establish and maintain social relations across geographical and institutional barriers. By the assumptions of the "death of distance" hypothesis, this removal of barriers to interaction opportunity should lead to a "flattening" of tie distributions with respect to distance. At the same time, however, it is not inevitable that the theoretical possibility of long-range interactions will translate to the loss of geography as a practical constraint on social ties. Geographical proximity continues to be associated with opportunities for common organizational affiliations, shared environmental exposures, and meeting opportunities; even if ties *can* be created and maintained at large distances, then, the contexts that initiate and sustain social relationships remain largely geographical. Whether distance is in fact "dead" as a determinant of large-scale social interaction in ostensibly barrier-free settings such as OSNs thus remains an important question, with ample reason (e.g., Backstrom et al. 2010) to doubt the "flat world" narrative.

In this work, we examine the question of whether and how physical geography and other basic opportunity structures continue to predict the structure of large-scale interpersonal networks. To provide a strong test of the ongoing importance of geography, we stack the deck *against* distance by deliberately choosing a social environment in which geographical opportunity structures are theoretically as weak as

possible: OSNs. As noted above, OSNs offer a context in which social ties can be formed with almost anyone, almost anywhere, and almost at any time. Moreover, social ties in OSNs can cost relatively little to create and maintain (with maintenance costs, if any, being unrelated to the distance between interaction partners). If traditional opportunity and cost constraints on social interaction are dead, this should be their graveyard. If one continues to find that factors such as geography, status, wealth, and other such factors play a strong role in the structure of online social ties, however, then there is good reason to believe that technological changes will not eliminate traditional determinants of social structure in other (less "barrier free") contexts.

To examine the association between social relationships and physical space, we use a large probability sample of publicly visible egocentric Facebook networks for which the individuals involved specify university affiliations (Gjoka et al. 2010). This is one of only a few data sets providing a probability sample of large-scale, spatially embedded, online social relationships with interesting individual-level social status differences. In addition, it offers a conservative context in which to study the influence of geography on social ties. Not only are friendship initiations relatively easy on Facebook, but individual users have access to a large, diverse (both spatially and socially) set of potential contacts. Indeed, the environment of Facebook itself is designed to lessen distance effects, precisely as the "death of distance" hypothesis alleges. Insofar as geography continues to matter in this setting, its importance in other (e.g., face-to-face) contexts is likely to be even *more* pronounced. We focus on university-based ties because literature in both higher education and social networks speaks to the importance of social ties and potential institutional status effects that may structure such relationships (e.g., as a mechanism to replicate inequality; see Rivera 2011). Thus, this case allows us to test the impact of nongeographical institutional factors on tie structure (e.g., social distance factors; Blau 1977; McPherson 2004).

To foreshadow our findings, we show here that distance is still a vital predictor of social ties, even in a case in which these effects would be, in theory, expected to be at their weakest. Our analyses demonstrate that the probability that any two individuals are tied drops significantly as the physical distance between them increases. In addition, there are clear differences in how distance operates that are organized along status lines. Specifically, we find that long-range ties are more frequently sustained between those affiliated with private universities than those affiliated with public institutions. Finally, our analyses indicate that institutional prestige differences also add to the effects of distance, with ties between affiliates of schools with disparate prestige being less common than those between affiliates of similarly prestigious institutions. Overall, these results suggest a story of continuing stratification along geographical and institutional lines: distance is not dead in the Internet age, and neither are other classical determinants of social structure. These continuing lines of division have important consequences for the structure of interpersonal networks and for the ongoing reproduction of inequality in settings (such as higher education) for which personal ties act as conduits for status-relevant resources.

Study Setting

The OSN Facebook offers a rich context in which to study social interaction. As such, it has attracted researchers from many different fields (Lewis et al. 2008; Tufekci 2008; Walther et al. 2008; Wimmer and Lewis 2010). Users of the site build detailed personal profiles including information on demographics, interests, and activities. Beyond personal characteristics, Facebook allows users to publicly declare "friendships" with other users, resulting in a massive online social network. Declared friendships must be confirmed by both parties involved to be realized, and therefore constitute mutual relationships acknowledged by both individuals. These relationships are viewed as socially meaningful by users of the system and have been found to afford users both benefits and consequences (Acquisti and Gross 2006; Ellison, Steinfield, and Lampe 2007; Mazer, Murphy, and Simonds 2007; Tufekci 2008) in addition to serving as conduits for information flow (including ongoing flows of information regarding personal events and activities that are by default shared automatically between declared friends).

Given its extremely high membership rates (more than 1 billion users worldwide, as of this writing), Facebook users have access to an extremely large, diverse (both spatially and demographically) population of potential social contacts.[2] Indeed, previous research has shown that traditional tie creation mechanisms such as racial homophily have less influence in this environment than in conventional settings (Wimmer and Lewis 2010). Combined with the fact that creation and maintenance of social ties is essentially costless, this theoretically infinite set of potential contacts results in an ideal case for equal mixing across traditionally stratified groups. Moreover, an individual's limited capacity to maintain social ties is effectively less restricted in this context than it would be in typical face-to-face social environments, because of the digital bookkeeping features of online social networks such as Facebook. In fact, removing social ties requires more effort on the part of the user than simply letting them remain (i.e., it is both a potentially stigmatized act and one that requires a specific action on the part of the user). In addition, the Facebook infrastructure "encourages" users of the site to form social relationship on the basis of shared interests and activities. For example, one can search for other users with identical listed favorite books, films, or music.

Many scholars have explored the social norms of Facebook friending, as well as relationship between Facebook friendship and other types of social relationships (e.g., see Ellison, Steinfield, and Lampe 2006; Lampe, Ellison, and Steinfield 2006). Although our focus is on

structure rather than norms, we note that much of this work supports the empirical findings presented here: traditional opportunity structures still matter. The most important feature of the Facebook environment for our research per se is the fact that any costs associated with the creation, maintenance, and/or deletion of social ties do not depend explicitly on physical distance or academic institution. Thus, the social environment on Facebook is ideal for lessening the effects of physical distance, as well as other barriers to social relationships. Any systematic differences in mixing between groups present in this context will likely be further exaggerated in face-to-face settings—the persistence of interaction barriers within Facebook is thus a strong indicator for their continuing relevance in other social arenas.

Facebook Data

The specific data used in this research come from a uniform sample of Facebook users collected by Gjoka et al. (2010). This data set has been used to explore many methodological and social phenomena in the computer science and social science literatures (e.g., Almquist 2012; Gjoka, Butts et al. 2011; Gjoka, Kurant et al. 2011; Kurant et al. 2011, 2012).

These data were collected by sampling directly from the population of 32-bit user identification numbers. We use a rejection sampling procedure to guarantee a truly uniform sample of users from the existing space of all publicly shared Facebook profiles (Gjoka et al. 2010).[3] It is one of the largest principled samples of the Facebook social network in existence. Importantly for our purposes, it does not suffer the restriction of being constrained to a few universities or colleges, nor does it suffer from biases due to nonprobability sampling (e.g., via breadth-first search), as do many other such data sets (e.g., Lee, Scherngell, and Barber 2011; Lewis et al. 2008; Wimmer and Lewis 2010). From these data we are able to obtain a probability sample of nodes and edges from the large-scale, spatially embedded social network of friendship ties between university-affiliated individuals.

Our sample consists of approximately 1 million users chosen uniformly at random from the publicly visible population of Facebook. For each of these users, we have a list of all public friendship ties between ego and his or her alters. When specified, we also obtain university affiliation.[4] With these data, it is possible to count the number of observed ties that exist between individuals at each pair of institutions. That is, given two schools A and B, we are able to observe the number of friendship ties between individuals affiliated with each school (as well as the number of users with each affiliation). Table 1 provides an illustrative example of the data in question, showing the observed counts of ties between four top-ranked universities as well as the size of the sampled user population. We use these sample quantities to model tie probability between users, as described below.

Table 1. Example of Social Ties between Top U.S. Universities.

	Harvard	Princeton	Yale	Stanford
Harvard	5,420	690	1,349	2,002
Princeton	690	2,740	589	353
Yale	1,349	589	4,220	459
Stanford	2,002	353	459	4,214
Estimated user population size	1,408	2,383	2,765	2,469

Geography, Institutional Context, and Status

Traditional studies of propinquity reveal that social interaction may be significantly affected by physical distance, even at very small scales; physical barriers such as the location of housing units, the orientation of walkways, or the location of desks in an office have been shown to affect interaction propensity (Barnlund and Harland 1963). Despite extensive historical evidence supporting the importance of physical distance as a determinant in social relationships, the continued importance of such factors has been questioned in the current environment of rapid technological change. Evaluating such claims with the case of online friendships offers a conservative test of the "death of distance" hypothesis. To do so, we require spatially embedded social ties. Restricting the set of social relationships of interest to those among university-affiliated persons allows us to associate a physical location with each individual in our data set. Each academic institution was geocoded on the basis of the published latitude and longitude coordinates in the Google Maps application programming interface (https://developers.google.com/maps/?csw=1).

Beyond physical distance, this choice of data was also motivated by questions about institutional opportunity structures. Ample evidence exists demonstrating the effect of institutional or status characteristics on social relationships (e.g., Currarini, Jackson, and Pin 2010); however, less work has examined how these institutional factors moderate the relationship between distance and social interaction. Because of these known mechanisms, it is important to control for nongeographical effects expected within these data. To control for the impact of nongeographical institutional factors on tie structure, we collect data on institutional covariates such as perceived status, size, setting, and so forth. Institutional prestige in the context of universities can be viewed as a proxy for access to resources, reputation, and cultural presence (i.e., likelihood that a person has heard of a given university), each of which could potentially influence the probability of ties between schools. For example, higher prestige universities have larger endowments (Zemsky 2003), which for example translates into access to resources for students to intern in faraway cities, attend national competitions, or

participate in overseas exchanges (Zemsky 2003). Furthermore, institutions of higher prestige recruit from larger regions both domestically and internationally than do lower prestige universities (e.g., Clark 2004). Moreover, these effects of prestige could be further exacerbated by the different levels of media attention and salience of the higher prestige universities, with members of high-status universities attracting more attention (ceteris paribus) than members of low-status universities.

Measuring institutional status within the higher education system is a complex issue; in this work we are interested primarily in public awareness and perceived status, rather than measures of quality or value. Thus, we use highly publicized national university rankings issued by *U.S. News & World Report* (*USNWR*) (McDonough et al. 1998), a well-known and influential source of ratings during the study period.[5]

Although *USNWR* rankings contain 262 national academic institutions, we limit our study to their tier 1 and tier 2 institutions because these are the only ones that are both ranked and scored.[6] This constitutes a total of 196 top academic institutions during the study period, ranging from Harvard University to Andrews University. For each of these 196 schools, we have a set of institutional-level covariates, such as the number of undergraduates, endowment, type of academic calendar, religious affiliation, tuition, setting, and so on. These institutional statistics are released with the *USNWR* rankings.

Methods

A Note on Social Network Concepts

Throughout this work we use basic concepts and notation from social network analysis. A social network consists of a set of entities, together with a relation on those entities (see Wasserman and Faust 1994). The set of potential relations that might occur on the set of entities in a social network is extremely varied; relations could constitute marriage, communication, association, copresence, or other forms of social interaction. For our purposes we require that relations be defined on pairs of actors (here on referred to as dyads). Specifically, the relationship of interest here consists of mutually acknowledged friendship ties between pairs of individuals. We represent social networks formally as *graphs*. (Following common practice, we use the terms *network* and *graph* interchangeably throughout our subsequent discussion.) A graph is a relational structure consisting of a set of entities (called vertices or nodes), and a set of connections among pairs of entities (called edges or ties). Formally, we represent a graph by the pair $G = (V, E)$, where V is the vertex set and E is the edge set. These basic network concepts will aid in our discussion of spatial networks and the methods used in this research.

Spatial Bernoulli Graphs

To examine the relationship between spatial (and other) factors and social ties, we use a scalable family of nonlinear statistical models previously described (Almquist and Butts 2012; Butts and Acton 2011; Butts et al. 2012). These models are closely related to the gravity models of the spatial econometrics literature (Haynes and Fortheringham 1984), and are a special case of the general exponential random graph framework (see Wasserman and Robins 2005 for a review). We represent spatial influences by treating individuals as being associated with particular points in space; the existence (or nonexistence) of ties is then assumed to arise from a discrete exponential family conditional on the realized interpoint distances (Butts and Acton 2011). One can express this family of models as follows:

$$\Pr(\mathbf{Y} = \mathbf{y} | \mathbf{D}) = \prod_{\{i,j\}} B(Y_{ij} = y_{ij} | \mathcal{F}_d(D_{ij})), \qquad (1)$$

where $\mathbf{Y} \in \{0,1\}^{N \times N}$, $\mathbf{D} \in [0, \infty)^{N \times N}$, $\mathcal{F}_d : [0, \infty) \mapsto [0,1]$, and B is the Bernoulli pmf. Y here represents the graph adjacency matrix (with observed value $\mathbf{y}$), such that the (i, j) cell in the matrix is equal to 1 if there exists a tie from node i to node j and 0 otherwise. D is a matrix of distances between nodes, such that the (i, j) cell represents the distance from node i to node j. The *spatial interaction function* (SIF), $\mathcal{F}_d$, is a core component of this modeling framework; it relates distance to the marginal tie probability. Indeed, model selection will depend largely on choosing the form and components of the SIF.

Spatial graph models exploit the observation that marginal tie probability in social networks generally changes (typically decreasing) as distance between individuals increases. This basic property is supported by empirical findings from many different fields (Almquist and Butts 2015; Boessen et al. 2014; Bossard 1932; Festinger et al. 1950; Freeman, Freeman, and Michaelson 1988; Hägerstrand 1967; Latané, Nowak, and Liu 1994; McPherson, Smith-Lovin, and Cook 2001; Smith et al. 2015). Although the model of equation 1 treats edges as conditionally independent given distance, its large-scale behavior is robust to unmodeled dependence under a range of conditions (e.g., Butts 2003, 2011). This is particularly true given that we are here interested in aggregate tie volumes between university-affiliated groups (an extremely robust property) rather than the detailed structure of networks at the micro level (e.g., centrality scores or local clustering). Despite their simplicity, spatial Bernoulli graphs have been demonstrated to predict social processes ranging from regional identification (Almquist and Butts 2015) to crime rates (Hipp et al. 2013).

The SIF may take many different functional forms, each of which has important theoretical implications for the macroscopic properties of the networks it generates (Butts 2010). Although one can choose functional forms in an exploratory

manner, knowledge of the relational setting can also guide the choice of SIF. One must take account of the context of interaction, which may in fact rule out particular function forms (for a detailed description, see Butts 2003; Butts and Acton 2011). From a hypothesis-testing perspective, model selection that considers different function forms rejects specific relationships (or theories) between distance and tie probability, along with their corresponding implications for network structure. For illustrative purposes, consider a simple power law functional form (Butts and Acton 2011). In this case, tie probability decays as a power law in distance:

$$\mathcal{F}_d(x) = \frac{p_b}{(1+\alpha x)^{\gamma}}, \qquad (2)$$

where $0 \leq p_b \leq 1$ is a baseline tie probability, $\alpha \geq 0$ is a scaling parameter, and $\gamma > 0$ is the exponent that controls the distance effect. The power law function is monotone decreasing in distance, which implies that edges on average tend to be short; adjacent individuals are closer in space than nonadjacent individuals, on average. This type of relationship tends to produce local clustering. Power law functions also have relatively heavy tails. Heavy-tailed interaction functions tend to produce networks in which there are relatively large numbers of long-range ties, thus creating short paths between spatially distant regions in the network. When combined with the heterogeneous distribution of population through space, SIFs such as that of equation 2 can produce surprisingly complex network structures (see, e.g., Butts et al. 2012). Here, we further extend this model to allow the SIF itself to vary as a function of other covariates, as described below.

Spatial Bernoullli Graphs with Covariates

Distance is known to be an extremely powerful determinant of social interaction, as previous research has documented (Latané et al. 1995; Stewart 1941b). However, it is not clear that the influence of distance will be constant across pairs of individuals with varying social or institutional contexts. In particular, we are interested in the differential effects of distance on tie probability along status lines, that is, how the marginal relationship between distance and tie probability might vary across different stratification categories and social groups. To incorporate the possibility of inhomogeneous distance effects, we add covariate terms into the spatial Bernoulli model. This extends the model family in a simple manner, allowing one to capture the impact of tie-level covariates such as difference in institutional prestige on the SIF. Introducing a linear predictor for each of the SIF parameters allows one to hypothesize about the differential influences of tie variables on the probability of a social connection in terms of the base tie probability, scale, or shape of the distance–tie probability relationship. In this framework, the general SIF form for the power law model described above becomes

$$Pr(Y_{ij} = 1) = \mathcal{F}_d(D_{ij}, p_{b_{ij}}, \alpha_{ij}, \gamma_{ij}) = \frac{p_{b_{ij}}}{(1 + \alpha_{ij} \cdot D_{ij})^{\gamma_{ij}}}, \qquad (3)$$

where $p_{b_{ij}} = \mathrm{logit}^{-1}(\theta^T X_{ij})$, $\alpha_{ij} = \exp(\psi^T W_{ij})$, $\gamma_{ij} = \exp(\phi^T U_{ij})$, θ, ψ, and ϕ are parameter vectors, and X, W, and U are covariate matrices.

Covariates may enter into the model at different points, each of which has specific effects on the subsequent relationship between distance and tie probability. It is important to recognize that (as for all network models) one must frame each effect in terms of tie-level covariates. As such, the characteristics of a given individual himself or herself cannot alone determine tie probability; rather, it is the attributes of the *pair* that matter. For example, we might hypothesize that ties between two men (or women) are more likely than ties in cross-gender relationships at all distances. Including a covariate for whether ties are homogeneous or heterogeneous by gender would capture this type of effect.

Covariate effects interact with tie probability via each of the three model parameters, p_b, α, and γ in distinct ways. For instance, as mentioned previously, we might hypothesize gender homophily to be important in determining tie probability. One possibility is that homophily acts as a constant factor across distance, systematically boosting the relative probability of gender-homophilous versus gender-heterophilous ties. In this case, gender consonance would occur as a covariate effect in the baseline probability parameter, p_b. (Effects of this general sort have been considered, e.g., by Daraganova et al. 2012.) On the other hand, the impact of this or other covariates might indeed differ across distance. Socioeconomic status differences, for example, are likely to be exaggerated at long distances. Higher socioeconomic status classes have more resources and potential for mobility, making distance less constraining on social interaction. These types of effects are then likely to influence α or γ, rather than p_b. As a general matter, p_b effects adjust the *base probability* of a tie between two individuals, regardless of spatial position. α effects change the *effective scale* on which distance operates, making a given physical unit of distance "shorter" or "longer" in terms of its impact on a given pair of individuals. Finally, γ effects change the *shape* of the distance–tie probability relationship; for power law models, increasing γ produces SIFs with heavier tails, allowing a greater chance of ties between individuals at very long distances. A given covariate (e.g., concordance on gender) may affect one of these parameters, or it may do so in combination. In the latter case, the nonlinear interaction between the base probability, scale, and shape effects will determine the overall impact of the covariate on the distance–tie probability relationship. Extending spatial Bernoulli models to include tie-level covariates this allows one to explore a wide range of hypotheses about the differential effects of distance on tie probability.

Model Fitting and Selection

Although the spatial Bernoulli modeling framework described above is relatively straightforward in theory, the estimation procedures required to fit spatial Bernoulli models with tie-level covariates to large-scale, real-world data are nontrivial. We use custom software to fit these models to data, which incorporates the statnet software suite (Handcock et al. 2008) for use in calculating statistics on the egocentric networks. We consider four basic forms for the SIF—exponential, logistic, power law, and attenuated power law (Butts 2003). Along with these four SIF functional forms, we explore a series of institutional-level covariates, including the difference in endowments of the schools with which the pair of individuals are affiliated, the difference in tuition, and the difference in setting of the schools (e.g., urban or rural). Each of these potential predictions begin as individual-level covariates and must be transformed into node-pair covariates. Thus we consider the absolute difference, sum, and mean of each of these factors. The interpretation in this case would be that the absolute difference in endowment between two schools could influence the chance that individuals at these school are friends, for example.

Measures of university status are of particular interest, because they speak to the ways in which distance may differentially affect social interaction across social categories, that is, the interaction between physical and social distance. One of the most basic distinguishing factors among academic institutions is whether they are public or private. As such we classify ties on the basis of the type of institution at each end point. This results in a three-fold classification of social ties.

The *USNWR* data allow the construction of a "prestige" measure for each of the top-ranked universities. The methodology of the rankings is somewhat obscure; by some combination of factors, each school ultimately receives a numerical scoring, by which it is later ranked. Rank and score are also highly correlated with selectivity (determined primarily by acceptance rates). Selectivity is a four-category distinction ranging from most to least selective. Together these three factors are a suitable proxy for the prestige of a university. Unsurprisingly, these three factors are highly correlated. Therefore, we take *prestige* to be the first principal component scores of these three covariates. The first principal component captures just over 99 percent of the variance in these factors.

In addition to the primary covariates of interest, public or private category, and prestige differences, we also consider differences in endowment, tuition, and setting. For numerical covariates, we consider absolute difference, mean, and median, each of which offers a potential relationship between status indicators and the marginal tie probability. Given the set of possible SIF functional forms and covariates, model selection in this context presents a formidable task. We thus use an automated procedure to find likely candidate models and compare models via goodness-of-fit criteria. The procedure of model selection can itself be viewed as a means of rejecting (or failing to reject) hypotheses about the form of the SIF and of the influence of different covariates. Each of the potential explanatory factors not selected, given some selection criteria, can be seen as failing to reject the null hypothesis about the influence of this predictor, essentially indicating that this set of covariates bears no significant relationship to the quantity of interest, the probability of a social tie.

We evaluate each of the models using a penalized deviance measure that takes into account the issue of model overfitting; in this case we have opted to use the Bayesian information criterion (Schwarz 1978), which is the generally preferred model selection criterion for exponential family models. The top six models can be seen in Table 2, along with an intercept-only model for comparison. Once the top model is determined, we estimate the model parameters via posterior maximization, as shown in Table 3.[7]

Again we stress that the covariates found in Table 2 were not the only ones at risk for inclusion in the model—nonselection of a covariate can thus be interpreted as a de facto rejection of the hypothesis that it is a net predictor of tie probability given included effects. The absolute difference in prestige and indicators for whether the schools involved are private or public seem to be the most powerful explanatory factors in this case. They appear in all of the top models in various forms. We also find that power law forms of the SIF perform best.

Results

Distance Still Matters

Let us begin by considering the overall aggregate relationship between distance and online friendship ties on Facebook. Although new technologies have undoubtedly changed the interplay between social interaction and distance, it is unclear that they have totally obliterated all effects of distance. Traditional theories predict that tie probabilities decrease as distance increases, that is individuals who are closer in physical space are more likely to be tied than those who are far apart. As noted above, some scholars have argued that the world is now "flat" and that tie probabilities no longer decrease with distance.

Consider a simple plot of the raw mixing rates in the data as a function of distance. Recall these rates are calculated by considering the proportion of friendship ties that are observed in relation to those that could have been observed (i.e., the potential number of ties on the basis of the estimated population of the two different schools). Figure 1 shows a clear relationship between distance and tie weight in the observed friendship network among individuals on Facebook affiliated with U.S. universities. We also show a visualization of these raw mixing rates in Figure 2. It is evident that the chance of a tie decreases

Table 2. Model Selection Results for the Facebook Friendship Network.

Model	p_b Effects			α Effects			γ Effects			Spatial Interaction Function Form	BIC
Covariate	Intercept	Public/Private	Prestige	Intercept	Public/Private	Prestige	Intercept	Public/Private	Prestige		
Model 1	✓	✓	✓	✓	✓	✓	✓	✓		Power law	24,911,904
Model 2	✓	✓	✓	✓	✓		✓	✓	✓	Power law	24,918,710
Model 3	✓	✓	✓	✓	✓		✓	✓		Attenuated power law	24,926,060
Model 4	✓	✓	✓	✓		✓	✓	✓	✓	Attenuated power law	24,933,741
Model 5	✓		✓	✓	✓	✓	✓	✓		Attenuated power law	24,935,807
Model 6	✓			✓			✓			Attenuated power law	25,139,114

Table 3. Parameter Estimates for Spatial Embedded Network Model of Friendship between Nationally Ranked U.S. Universities.

Parameter	Component	Estimate	Posterior Standard Deviations
p_b	Intercept	−6.0974	0.0061[a]
	Private-public	−0.4340	0.0200[a]
	Public-public	−0.7501	0.0063[a]
	Prestige	−0.0176	0.0000[a]
α	Intercept	2.1687	0.0259[a]
	Private-public	−2.2169	0.0493[a]
	Public-public	−4.5387	0.0269[a]
	Prestige	−0.0187	0.0001[a]
γ	Intercept	−1.0789	0.0016[a]
	Private-public	0.4523	0.0026[a]
	Public-public	1.0009	0.0023[a]

[a]All estimates >> 3 posterior standard deviations from zero.

with increased distance between contacts and is a least weakly monotonic as the logarithm of the distance. Thus, this fits with previous research in the area suggesting an inverse relationship. Is distance dead? No. Even in the case of online social connections, distance remains a powerful determinant of interaction probability. We explore the precise form of this relationship in greater detail presently.

The Public-private Divide

Table 3 contains parameter estimates and associated posterior standard deviations for the power law model (chosen in the previously described model selection process). The table indicates different predictors for each of the model parameters: the baseline tie probability (p_b), scaling factor (α), and exponent (γ) for the power law model. Contrary to what might be expected, the impact of these covariates is not confined to base tie probability, distance scale, or interaction shape; instead, they jointly alter the SIF in multiple, simultaneous ways. The nonlinearity of the this model makes individual parameter estimates difficult to interpret on their own. We thus find it more illuminating to consider the predicted tie probability curves at different representative values of the predictors in relation to distance.

The predicted values for the observed data are seen in Figure 3. Two important results stand out in this figure. First, we observe a banding in the tie probabilities by whether the relationship is between individuals at two public schools, two private schools, or one of each. Two individuals are more likely to be tied if they are at two private schools than at two public schools, at the same distance apart; this relationship seems to hold across all distances. Additionally, we find that the variance of the private-private tie probabilities is larger than that of the public-public schools, as seen in the wider band. Public-private schools are somewhere in the middle. Together, these results suggest that individuals at private schools are more likely to have social connections on Facebook that span larger distances. However, we also find that (because of prestige effects) persons affiliated with private schools are likely to exhibit more variability in their tie probabilities.

Consider the predicted curves for an average prestige difference in each of the three categories of tie classification by school type: public-public, public-private, and private-private. These results are seen in the center panel of Figure 4. We find significant differences in the tail weights of these curves. The tie probability between two private university–affiliated persons is much higher than two public university–affiliated persons or even one private and one public at large distances. This indicates that although distance itself has a significant effect on the chance of a social tie, this effect is exaggerated for ties involving public universities. Social ties between those affiliated with private academic institutions are found at higher rates across longer distances. If we consider two private school and two public school affiliated pairs, with each located at an identical, long distance from each other, the chance of a tie between the two private school affiliates is higher than the public school affiliates. This public-private status characteristic of the institutions to which people belong moderates the effect of distance on the chance of a social relationship.

Status Effects

Figure 4 shows the fitted prediction curves for the tie probability as a function of distance if we vary the prestige difference between persons' institutions. We show the 25th, 50th,
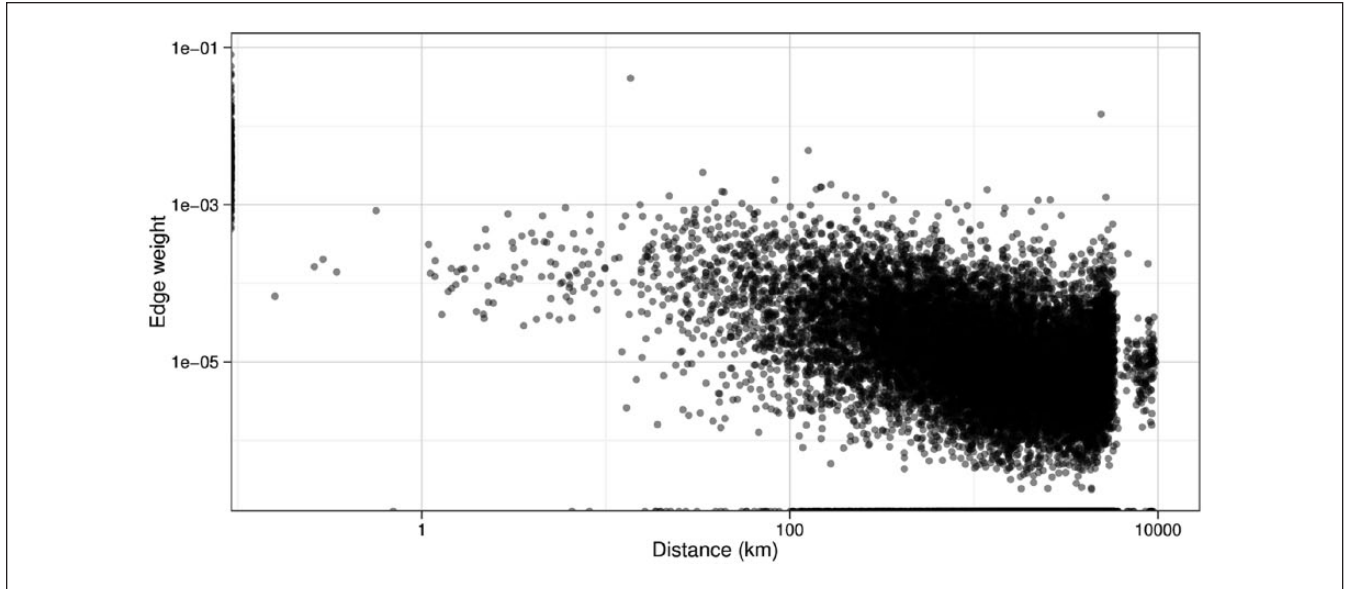
Figure 1. Effects of distance on tie frequency in observed data. Tie frequency decreases as a function of distance between end points.
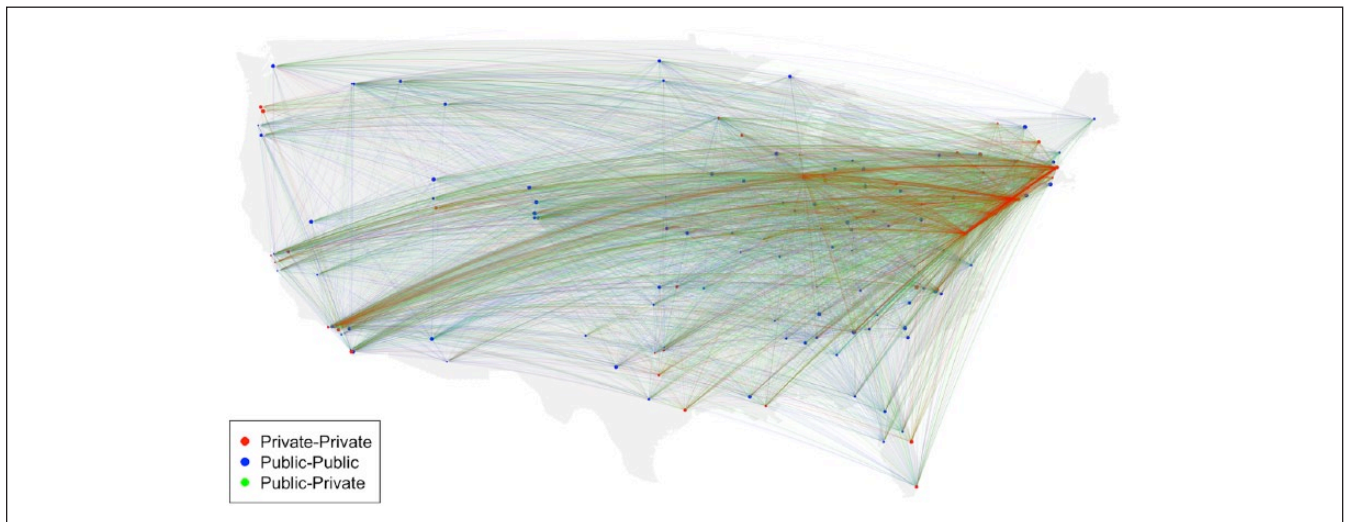

Figure 2. Social ties among U.S. universities.

and 75th percentiles of the prestige difference distribution. Evident in this figure is the result that larger differences in prestige *decrease* the probability of a friendship tie across all distances; the curves shift down as prestige differences become larger. Individuals at universities that differ greatly in their perceived prestige are less likely to be tied than those at universities of similar prestige. These results point toward a stratification of ties along lines of university prestige. Again we find that the relationship between distance and social ties is structured along status lines, producing systematic differences in the natures of an individual's social network that are organized by status categories.

Another feature of these curves is the crossing: the point at which the tie probability between a pair at two public

schools surpasses the tie probability between a pair at two private schools or even one private and one public school. This phenomenon occurs between distances of approximately 3 and 300 km; note that this distance range will often be within a state boundary. Thus, individuals affiliated with public schools in a given state are more likely to be tied to others associated with public schools in the same region at higher rates than individuals at two private schools in the same region. We discuss some of the potential mechanisms for producing this result below.

Together, these results offer a new perspective into the systematic differences in social network structure among university-affiliated persons on Facebook. Differences seem to exist along status lines, with an overarching global impact
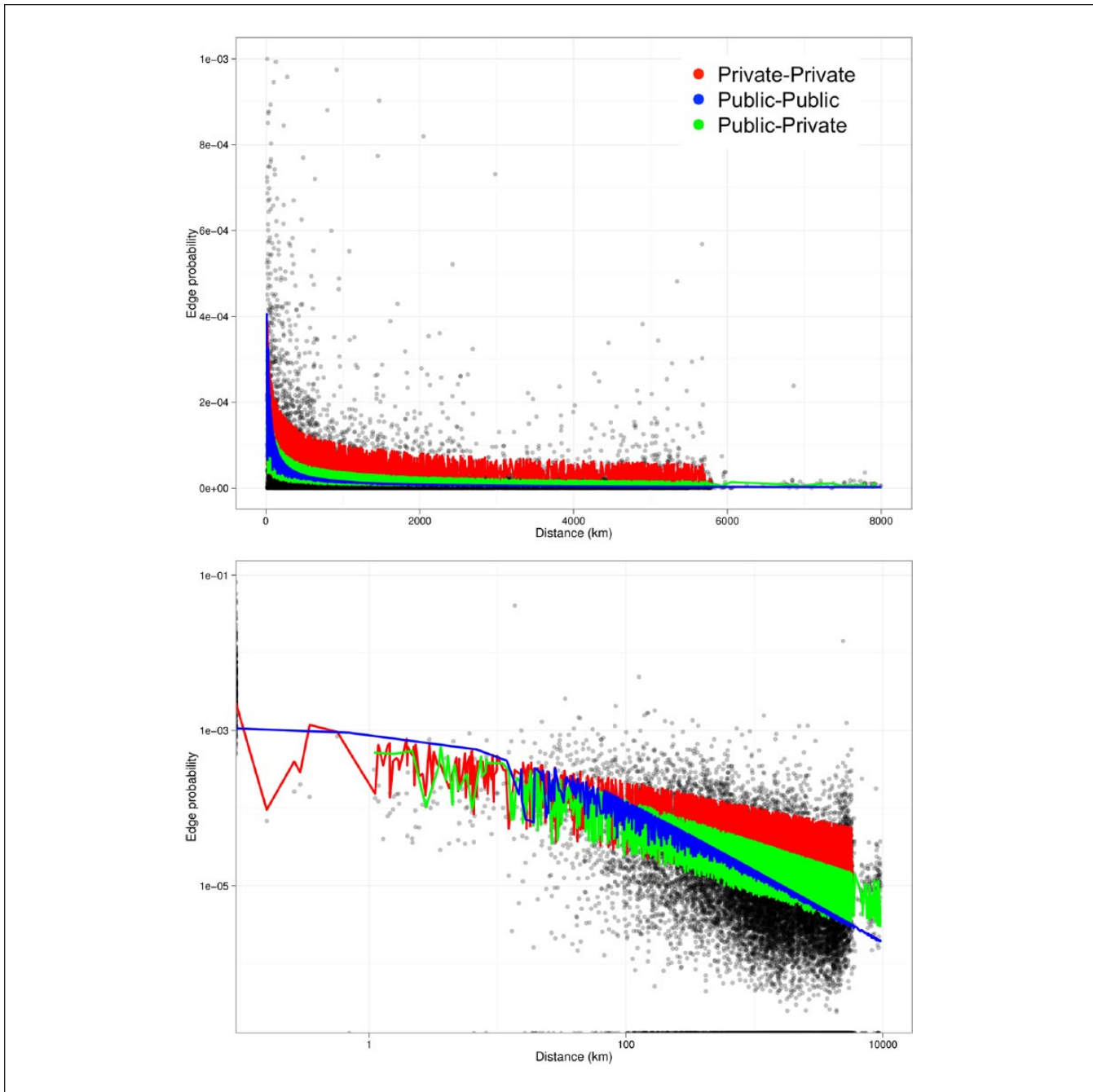
Figure 3. Fitted values of the spatial Bernoulli model. Ties between two private school are in red, two public schools are in blue, and one public one private are in green. Shown on log-log scale on right. Note the banding by tie type and the difference in band widths, indicating difference in variability in tie probability.

of geographical proximity. We discuss the implications of these results in the subsequent section.

Discussion

Our results show that there is a continuing impact of spatial proximity on online relationships among university affiliates, with differences based on the nature of the universities in question; these differences seem to be structured along institutional and status lines. We have demonstrated that distance is still a vital predictor of social ties, even in online environments in which these effects would be expected to be at their weakest. Although long-range friendships certainly exist, the probability that any two individuals are tied drops significantly as the physical distance between them increases. This is consonant with previous work suggesting that online ties (like those among Facebook users) are not formed exclusively in a spatial environment but in fact reflect social
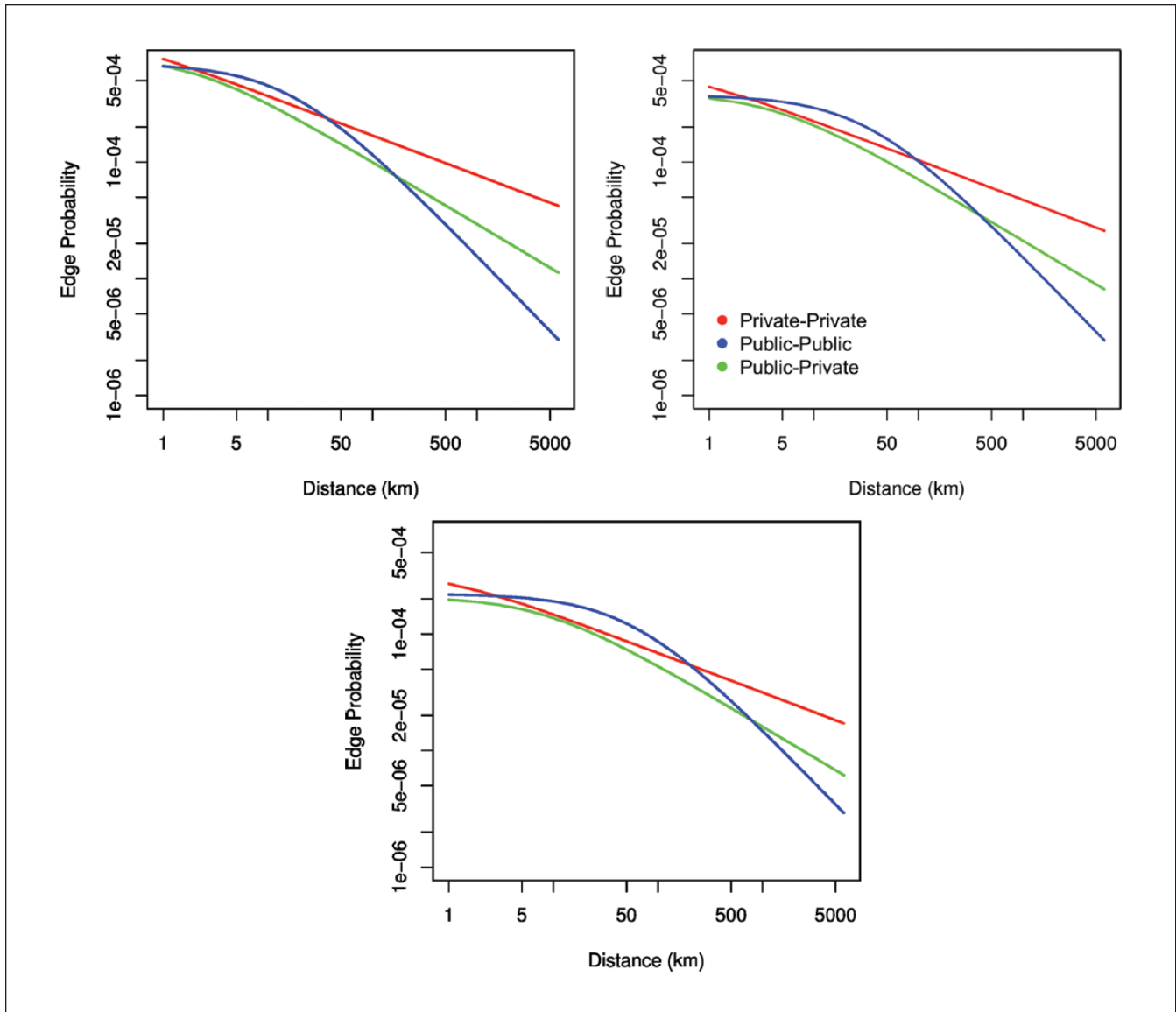
Figure 4. Impact of prestige difference and public-private differences on tie probability on log-log scale. The three figures show estimated spatial interaction functions at three different quantiles of prestige difference; as can be seen, tail weights and overall shape vary as a function of prestige.

processes that include offline and institutionally mediated dimensions. Although Facebook users could in principle form ties without regard to geographical limitations, they do not: they are disproportionately connected to those to whom they are more proximate.

Although propinquity is an overwhelming force, it does not affect everyone equally. We see clear differences between individuals at private versus public schools in the presence of long-range ties. Individuals affiliated with public schools are more "regional" in their social relations, and those individuals affiliated with private schools are more cosmopolitan. Because of the cross-sectional nature of this research, we were unfortunately not able to investigate the proportion of social ties that are formed before and after college attendance (e.g., see Noel and Nyhan 2011). This type of work would

offer additional insight into the mechanisms that lead to the observed differentiated social structure; for example, is it that individuals are primarily characterized by the social ties they made in high school, made in college, or a combination of these two processes? Our results may be the product of a migration process, a retention process, or both. Although this does offer an exciting avenue for future work, it is not the central concern of this research. Furthermore, our findings suggest that elites—those who attend top-tier private schools—show more dispersion in their friendship ties. It is conceivable that this is the result of an underlying selective migration process, in which elites preferentially scatter across the country ("pulling" old ties with them). Individuals who attend lower tier schools, on the other hand, may remain more regionally confined. Alternatively, the story might be

one of access. Some scholars have argued that the benefit of attending top-tier universities is the access it grants to the elite social networks of students and alumni (Dale and Krueger 2002; Gould 1989; Hoxby 2001; Hoxby and Terry 1999). Whether due to migration or due to the ability to tap into others' long-range contacts, the relative dispersion of ties at elite institutions may give those within them an advantage in tapping into a more diverse range of information and resources than those confined to more parochial environments (in close analogy to Burt 1992).

Finally, we explore the effects of prestige differences on social ties. Our findings indicate that prestige differences moderate the realized effects of distance. Higher differences in prestige between two schools lower the chance that social ties exist between affiliated individuals. Overall, these two results together suggest a story of stratification. Systematic differences in the properties of personal networks structured along status lines have important consequences for theories of inequality in higher education.

A college education is tied to a number of important life outcomes, such as employment, lifetime earnings, and class mobility; it is also tied to one's position within social structure and subsequent access to social capital (Granovetter 1974). Theories of higher education have always been concerned with the inequality in access and success in educational institutions (for recent work, see Bound, Hershbein, and Long 2009). There are many different mechanisms that may or may not be important to an individual's success within educational institutions and subsequent life outcomes; one important mechanism is that of access to resources via social ties (e.g., obtaining a job; Granovetter 1974). Although the advantages of attending a college or university are well known (e.g., lower unemployment, higher average earnings; Baum and Ma 2007), the additional benefit of attending highly prestigious private universities (e.g., Harvard University, Stanford University) compared with prestigious public schools (e.g., state flagship schools such as the University of California, Berkeley, or the University of Michigan) is in general contentious (although there is some evidence that attending more prestigious schools correlates with attending and acquiring admittance to prestigious graduate schools; Eide, Brewer, and Ehrenberg 1998). As Granovetter (1973) suggested, there are a number of reasons why the larger variation and distance structure of social ties might be advantageous. Individuals with more spatially diverse networks will have access to more varied resources and information; their networks will have the potential to ensure against local economic and disruptive conditions (e.g., local housing slumps, job loss, floods, fires).

There is a significant body of research on the effects of formal structure on the individual (e.g., behavior and environment; for a more detailed discussion, see Hurtado 2007). This research has been extended to help explain and interpret rates of college attendance, acceptance, and graduation at selective universities, especially in the context of segregation and racial discrimination (Espenshade 2009; Massey et al.

2006). This work has focused on a number of different theories, such as social integration and peer influence, as well as achievement measurements (Advanced Placement classes, SAT scores, grade point average, etc.) to explain who ends up at what university and who succeeds. However, few studies have focused on the explicit modeling of social interaction effects as a result of the university one chooses to attend. Here we are able to show that systematic difference do exist between individuals at top-tier universities; differences that are structure along status lines. We demonstrate explicitly that university affiliation is associated with inequalities in terms of network properties.

Conclusion

In this article we have examined the effects of physical distance and institutional indicators of prestige on social ties in a spatially embedded friendship network among top-tier U.S. universities by modeling the marginal relationship between distance and tie probability. We used data gathered from the OSN Facebook to construct a weighted friendship network among the 196 national universities ranked highest by *USNWR*. By locating these institutions in space, we are able to jointly explore the structure of the social network and the effects on this structure of space itself as well as external factors. As the results indicate, our findings reinforce the empirical evidence that distance structures many different kinds of social interaction and subsequent social ties by opportunities for tie formation. We have demonstrated that the probability of a friendship tie significantly decreases as distance between the entities involved increases, even in an online setting for which such effects would be expected to be at their weakest.

We find that there is a difference between the likelihood of two individuals' being tied on the basis of whether those individuals are affiliated with public or private schools and that this difference itself varies with geographical distance. This effect is further exacerbated by the prestige levels of the universities in question (i.e., two individuals are more likely to be tied if they both attend universities of about the same prestige level). Taken together, these effects demonstrate that universities help facilitate stratification not only in the traditional, offline context but also in the online environment. The lines of division known to sociologists from the past century continue to persist, even in the face of a changing technological environment.

Notes

1. Nor is this a new finding in the context of telephony; Hägerstrand (1967) famously showed similar results for phone calls in rural Sweden.

2. For the most recent statistics on U.S. online social networking statistics, see Pew Research Center (2012).
3. This sampling strategy was possible because of the unique nature of the Facebook user identification number space at the time of data collection; it is no longer feasible, making this data set extremely valuable.
4. University affiliation is a publicly visible characteristic that requires the user to have a verified university e-mail account. This is the only restriction on the user population; participants in the study could include students, faculty members, and staff members.
5. "In a relatively short period of time the U.S. News and World Report University Rankings annual ranking of the nation colleges and universities as undergraduate institutions has become the 'gold standard' of the ranking business" (Ehrenberg 2002).
6. *USNWR* does not rank the lower 25 percent of national universities, because of data limitations (U.S. News & World Report Online 2010).
7. All estimates reflect posterior modes under diffuse t priors; Monte Carlo procedures used in the estimation process were repeated 25 times, with the median estimate reported.

References

Acquisti, A., and R. Gross. 2006. "Imagined Communities: Awareness, Information Sharing, and Privacy on the Facebook." Pp. 36–58 in *Privacy Enhancing Technologies*, edited by George Danezis and Philippe Golle. New York: Springer.

Adams, P. C. 1997. "Cyberspace and Virtual Places." *Geographical Review* 87(2):155–71.

Almquist, Zack. 2012. "Random Errors in Egocentric Networks." *Social Networks* 34(4):493–505.

Almquist, Zack W., and Carter T. Butts. 2012. "Point Process Models for Household Distributions within Small Areal Units." *Demographic Research* 26:593–632.

Almquist, Zack W., and Carter T. Butts. 2015. "Predicting Regional Self-identification from Spatial Network Models." *Geographical Analysis* 47(1):50–72.

Backstrom, Lars, Eric Sun, and Cameron Marlow. 2010. "Find Me if You Can: Improving Geographical Prediction with Social and Spatial Proximity." Pp. 61–70 in *Proceedings of the 19th International Conference on World Wide Web*. New York: Association for Computing Machinery.

Barnlund, Dean C., and Carroll Harland. 1963. "Propinquity and Prestige as Determinants of Communication Networks." *Sociometry* 26(4):467–79.

Baum, Sandy, and Jennifer Ma. 2007. "Education Pays for Individuals and Society." *Higher Education* 82(1):1–48.

Benniger, J. R. 1986. *The Control Revolution*. Cambridge, MA: Harvard University Press.

Bijker, W., and J. Law. 1994. *Shaping Technology/Building Society: Studies in Sociotechnical Change*. Cambridge, MA: MIT Press.

Blau, Peter M. 1977. *Inequality and Heterogeneity*. New York: Free Press.

Boessen, Adam, John R. Hipp, Emily J. Smith, Carter T. Butts, Nicholas N. Nagle, and Zack Almquist. 2014. "Networks, Space, and Residents' Perception of Cohesion." *American Journal of Community Psychology* 53(3–4):447–61.

Bossard, J.H.S. 1932. "Residential Propinquity as a Factor in Marriage Selection." *American Journal of Sociology* 38(2):219–44.

Bound, John, Brad Hershbein, and Bridget Terry Long. 2009. "Playing the Admissions Game: Student Reactions to Increasing College Competition." NBER Working Paper No. 15272. Cambridge, MA: National Bureau of Economic Research.

Burt, Ron S. 1992. *Structural Holes: The Social Structure of Competition*. Cambridge, MA: Harvard University Press.

Butts, Carter T. 2003. "Predictability of Large-scale Spatially Embedded Networks." Pp. 313–23 in *Dynamic Social Network Modeling and Analysis: Workshop Summary and Papers*, edited by Ronald L. Breiger, Kathleen M. Carley, and P. Pattison. Washington, DC: National Academies Press.

Butts, Carter T. 2010. "Space and Structure: Models and Methods for Large-scale Interpersonal Networks." Unpublished manuscript.

Butts, Carter T. 2011. "Bernoulli Graph Bounds for General Random Graphs." *Sociological Methodology* 41(1):299–345.

Butts, Carter T., and Ryan M. Acton. 2011. "Spatial Modeling of Social Networks." In *The SAGE Handbook of GIS and Society Research*, edited by Timothy Nyerges, Helen Couclelis, and Robert McMaster. Thousand Oaks, CA: Sage.

Butts, Carter T., Ryan M. Acton, John R. Hipp, and Nicholas N. Nagle. 2012. "Geographical Variability and Network Structure." *Social Networks* 34(1):82–100.

Cairncross, Frances. 2001. *The Death of Distance: How the Communications Revolution Is Changing Our Lives*. Cambridge, MA: Harvard Business Press.

Caplow, Theodore, and Robert Forman. 1950. "Neighborhood Interaction in a Homogeneous Community." *American Sociological Review* 15:357–66.

Castells, M. 1996. *The Rise of the Network Society*. Oxford, UK: Blackwell.

Clark, Burton R. 2004. "Delineating the Character of the Entrepreneurial University." *Higher Education Policy* 17(4):355–70.

Currarini, Sergio, Matthew Jackson, and Paolo Pin. 2010. "Identifying the Roles of Race-based Choice and Chance in High School Friendship Network Formation." *Proceedings of the National Academy of Sciences* 107(11):4857–61.

Dale, Stacy B., and Alan B. Krueger. 2002. "Estimating the Payoff to Attending a More Selective College: An Application of Selection on Observables and Unobservables." *Quarterly Journal of Economics* 117(4):1491–1527.

Daraganova, Galina, Pip Pattison, Johan Koskinen, Bill Mitchell, Anthea Bill, Martin Watts, and Scott Baum. 2012. "Networks and Geography: Modelling Community Network Structures as the Outcome of Both Spatial and Network Processes." *Social Networks* 34(1):6–17.

Diamond, Jared M. 1999. *Guns, Germs, and Steel: The Fates of Human Societies*. New York: W. W. Norton.

Durkheim, Emile. 1893. *The Division of Labor in Society*. New York: Free Press.

Ehrenberg, Ronald G. 2002. "Reaching for the Brass Ring: The *U.S. News & World Report* Rankings and Competition." *Review of Higher Education* 26(2):145–62.

Eide, Eric, Dominic J. Brewer, and Ronald G. Ehrenberg. 1998. "Does It Pay to Attend an Elite Private College? Evidence on the Effects of Undergraduate College Quality on Graduate School Attendance." *Economics of Education Review* 17:371–76.

Ellison, N. B. 2007. "Social Network Sites: Definition, History, and Scholarship." *Journal of Computer-Mediated Communication* 13(1):210–30.

Ellison, Nicole, Charles Steinfield, and Cliff Lampe. 2006. "Spatially Bounded Online Social Networks and Social

Capital: The Role of Facebook." Presented at the Annual Conference of the International Communication Association. Retrieved February 8, 2016 (http://www.ucalgary.ca/files/stas341/Facebook_ICA_2006.pdf).

Ellison, N. B., C. Steinfield, and C. Lampe. 2007. "The Benefits of Facebook 'Friends': Social Capital and College Students' Use of Online Social Network Sites." *Journal of Computer-Mediated Communication* 12(4):1143–68.

Espenshade, Thomas J. 2009. *No Longer Separate, Not Yet Equal: Race and Class in Elite College Admission*. Princeton, NJ: Princeton University Press.

Festinger, Leon, S. Schachter, and K. Back. 1950. *Social Pressures in Informal Groups*. Stanford, CA: Stanford University Press.

Freeman, Linton C., Sue C. Freeman, and A. G. Michaelson. 1988. "On Human Social Intelligence." *Journal of Social and Biological Structure* 11(4):415–25.

Friedman, Thomas L. 2005. *The World Is Flat: A Brief History of the Twenty-first Century*. New York: Farrar, Straus.

Gjoka, Minas, Carter T. Butts, Maciej Kurant, and Athina Markopoulou. 2011a. "Multigraph Sampling of Online Social Networks." *IEEE Journal on Selected Areas of Communication on Measurement of Internet Topologies* 29(9):1893–1905.

Gjoka, Minas, Maciej Kurant, Carter T. Butts, and Athina Markopoulou. 2010. "Walking in Facebook: A Case Study of Unbiased Sampling of OSNs." Pp. 1–15 in 2010 *Proceedings IEEE INFOCOM*. Piscataway, NJ: IEEE.

Gjoka, Minas, Maciej Kurant, Carter T. Butts, and Athina Markopoulou. 2011b. "Practical Recommendations on Crawling Online Social Networks." *IEEE Journal on Selected Areas of Communication on Measurement of Internet Topologies* 29(9):1872–92.

Glaeser, Edward L., and Giacomo A. M. Ponzetto. 2010. "Did the Death of Distance Hurt Detroit and Help New York?" Pp. 303–38 in *Agglomeration Economics*, edited by Edward L. Glaeser. Chicago: University of Chicago Press.

Gould, Roger V. 1989. "Power and Social Structure in Community Elites." *Social Forces* 68(2):531–52.

Granovetter, Mark. 1973. "The Strength of Weak Ties." *American Journal of Sociology* 78:1369–80.

Granovetter, Mark S. 1974. *Getting a Job: A Study of Contacts and Careers*. Cambridge, MA: Harvard University Press.

Hägerstrand, T. 1967. *Innovation Diffusion as a Spatial Process*. Chicago: University of Chicago Press.

Handcock, Mark S., David R. Hunter, Carter T. Butts, Steven M. Goodreau, and Martina Morris. 2008. "statnet: Software Tools for the Representation, Visualization, Analysis and Simulation of Network Data." *Journal of Statistical Software* 24(1):1–11.

Haynes, Kingsley E., and A. Stewart Fortheringham. 1984. *Scientific Geography, Vol. 2: Gravity and Spatial Interaction Models*. Beverly Hills, CA: Sage.

Hipp, John R., Carter T. Butts, Ryan Acton, Nicholas N. Nagle, and Adam Boessen. 2013. "Extrapolative Simulation of Neighborhood Networks Based on Population Spatial Distribution: Do They Predict Crime?" *Social Networks* 35(4):614–25.

Hoxby, Caroline M. 2001. "The Return to Attending a More Selective College: 1960 to the Present." Chap. 2 in *Forum Futures: Exploring the Future of Higher Education, 2000 Papers*, edited by Maureen Devlin and Joel Meyerson. San Francisco, CA: Jossey-Bass.

Hoxby, Caroline M., and Bridget Terry. 1999. "Explaining Rising Income and Wage Inequality among the College Educated."

NBER Working Paper No. 6873. Cambridge, MA: National Bureau of Economic Research.

Hurtado, Sylvia. 2007. "The Study of College Impact." Pp. 53–93 in *Sociology of Higher Education: Contributions ad Their Contexts*, edited by Patricia J. Gumport. Baltimore, MD: John Hopkins University Press.

Kolko, J. 2000. "The Death of Cities? The Death of Distance? Evidence from the Geography of Commercial Internet Usage." Pp. 73–98 in *The Internet Upheaval: Raising Questions, Seeking Answers in Communications Policy*, edited by Ingo Vogelsang and Benjamin M. Compaine. Cambridge, MA: MIT Press.

Kulshrestha, Juhi, Farshad Kooti, Ashkan Nikravesh, and P. Krishna Gummadi. 2012. "Geographic Dissection of the Twitter Network." Retrieved February 8 (http://www.mpi-sws.org/~farshad/geographic_dissection.pdf).

Kurant, Maciej, Minas Gjoka, Carter T. Butts, and Athina Markopoulou. 2011. "Walking on a Graph with a Magnifying Glass: Stratified Sampling via Weighted Random Walks." Presented at ACM SIGMETRICS '11, San Jose, California. Retrieved February 8, 2016 (http://mkurant.com/publications/papers/SIGMETRICS_2011.pdf).

Kurant, Maciej, Minas Gjoka, Yan Wang, Zack W. Almquist, Carter T. Butts, and Athina Markopoulou. 2012. "Coarse-grained Topology Estimation via Graph Sampling." Presented at ACM SIGCOMM Workshop on Online Social Networks (WOSN) '12, Helsinki, Finland. Retrieved February 8, 2016 (http://mkurant.com/publications/papers/wosn02-kurant.pdf).

Lampe, Cliff, Nicole Ellison, and Charles Steinfield. 2006. "A Face(book) in the Crowd: Social Searching vs. Social Browsing." Pp. 167–70 in *Proceedings of the 2006 Conference on Computer Supported Cooperative Work*. New York: Association for Computing Machinery.

Latané, B., J. H. Liu, A. Nowak, M. Bonevento, and L. Zheng. 1995. "Distance Matters: Physical Space and Social Impact." *Personality and Social Psychology Bulletin* 21(8):795–805.

Latané, Bibb, A. Nowak, and James H. Liu. 1994. "Measuring Emergent Social Phenomena: Dynamism, Polarization, and Clustering as Order Parameters of Social Systems." *Behavioral Science* 39(1):1–24.

Lee, Conrad, Thomas Scherngell, and Michael J. Barber. 2011. "Investigating an Online Social Network Using Spatial Interaction Models." *Social Networks* 33(2):129–33.

Lewis, Kevin, Jason Kaufman, Marco Gonzalez, Andreas Wimmer, and Nicholas Christakis. 2008. "Tastes, Ties, and Time: A New Social Network Dataset Using Facebook.com." *Social Networks* 30(4):330–42.

Massey, Douglas S., Camille Z. Charles, Garvey F. Lundy, and Mary J. Fischer. 2006. *The Source of the River: The Social Origins of Freshmen at America's Selective Colleges and Universities*. Princeton, NJ: Princeton University Press.

Mayhew, Bruce H. 1984. "Chance and Necessity in Sociological Theory." *Journal of Mathematical Sociology* 9(4):305–39.

Mazer, J. P., R. E. Murphy, and C. J. Simonds. 2007. "I'll See You on 'Facebook': The Effects of Computer-mediated Teacher Self-disclosure on Student Motivation, Affective Learning, and Classroom Climate." *Communication Education* 56(1): 1–17.

McDonough, Patricia M., Anthony Lising Antonio, Marybeth Walpole, and Leonor Xâochitl P. Âerez. 1998. "College Rankings: Democratized College Knowledge for Whom?" *Research in Higher Education* 39(5):513–37.

McPherson, J. Miller. 2004. "A Blau Space Primer: Prolegomenon to an Ecology of Affiliations." *Industrial and Corporate Change* 13(1):263–80.

McPherson, J. Miller, Lynn Smith-Lovin, and James M. Cook. 2001. "Birds of a Feather: Homophily in Social Networks." *Annual Review of Sociology* 27:415–44.

Mesch, G. S., I. Talmud, and A. Quan-Haase. 2012. "Instant Messaging Social Networks: Individual, Relational, and Cultural Characteristics." *Journal of Social and Personal Relationships* 29(6):736–59.

Mok, D., B. Wellman, and J. A. Carrasco. 2010. "Does Distance Still Matter in the Age of the Internet?" *Urban Studies* 47(13):2743–83.

Noel, Hans, and Brendan Nyhan. 2011. "The 'Unfriending' Problem: The Consequences of Homophily in Friendship Retention for Causal Estimates of Social Influence." *Social Networks* 33(3):211–18.

O'Brien, R. 1992. *Global Financial Integration: The End of Geography*. New York: Council on Foreign Relations Press.

Onnela, J. P., S. Arbesman, M. C. González, A. L. Barabási, and N. A. Christakis. 2011. "Geographic Constraints on Social Network Groups." *PLoS One* 6:e16939.

Pew Research Center. (2012). "Social Networking Fact Sheet." Retrieved February 8, 2016 (http://pewinternet.org/Commentary/2012/March/Pew-Internet-Social-Networking-full-detail.aspx).

Pool, I. S. 1977. *The Social Impact of the Telephone*. Cambridge, MA: MIT Press.

Rainie, L., and B. Wellman. 2012. *Networked: The New Social Operating System*. Cambridge, MA: MIT Press.

Rheingold, H. 2000. *The Virtual Community: Homesteading on the Electronic Frontier*. Cambridge, MA: MIT Press.

Rivera, Lauren A. 2011. "Ivies, Extracurriculars, and Exclusion: Elite Employers' Use of Educational Credentials." *Research in Social Stratification and Mobility* 29(1):71–90.

Schwarz, Gideon E. 1978. "Estimating the Dimension of a Model." *Annals of Statistics* 6(2):461–64.

Smith, Emily J., Christopher S. Marcum, Adam Boessen, Zack W. Almquist, John R. Hipp, Nicholas N. Nagle, and Carter T. Butts. 2015. "The Relationship of Age to Personal Network Size, Relational Multiplexity, and Proximity to Alters in the Western United States." *Journals of Gerontology Series B: Psychological Sciences and Social Sciences* 70(1):91–99.

Stewart, J. Q. 1941a. "The 'Gravitation,' or Geographical Drawing Power, of a College." *Bulletin of the American Association of University Professors* 27(1):70–75.

Stewart, J. Q. 1941b. "An Inverse Distance Variation for Certain Social Influences." *Science* 93(2404):89–90.

Takhteyev, Yuri, Anatoliy Gruzd, and Barry Wellman. 2012. "Geography of Twitter networks." *Social Networks* 34(1):73–81.

Tufekci, Zeynep. 2008. "Grooming, Gossip, Facebook and MySpace." *Information, Communication & Society* 11(4):544–64.

U.S. News & World Report Online. 2010. "Best Colleges 2011: About the Rankings/Methodology" (http://www.usnews.com).

Vogelsang, Ingo, and Benjamin M. Compaine, eds. 2000. *The Internet Upheaval: Raising Questions, Seeking Answers in Communications Policy*. Cambridge, MA: MIT Press.

Walther, Joseph B., Brandon van der Heide, Sang-Yeon Kim, David Westerman, and Stephanie Tom Tong. 2008. "The Role of Friends' Appearance and Behavior on Evaluations of Individuals on Facebook: Are We Known by the Company We Keep?" *Human Communication Research* 34(1):28–49.

Wang, Y., P. Lai, and D. Sui. 2003. "Mapping the Internet Using GIS: The Death of Distance Hypothesis Revisited." *Journal of Geographical Systems* 5(4):381–405.

Wasserman, Stanley, and Katherine Faust. 1994. *Social Network Analysis: Methods and Applications*. Cambridge, UK: Cambridge University Press.

Wasserman, Stanley, and Garry Robins. 2005. "An Introduction to Random Graphs, Dependence Graphs, and \$p*\$." Pp. 192–214 in *Models and Methods in Social Network Analysis*, edited by Peter J Carrington, John Scott, and Stanley Wasserman. Cambridge, UK: Cambridge University Press.

Wellman, Barry, Jeffery Boase, and Wenhong Chen. 2002. "The Networked Nature of Community: Online and Offline." *IT and Society* 1(1):151–65.

White, Lynn T. 1964. *Medieval Technology and Social Change*. Oxford, UK: Oxford University Press.

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(2):583–642.

Zemsky, Robert. 2003. "In Pursuit of Prestige: Strategy and Competition in US Higher Education." *Journal of Higher Education* 74(4):474–76.

Zipf, George K. 1949. *Human Behavior and the Principle of Least Effort*. New York: Hafner.

Author Biographies

Emma S. Spiro is an assistant professor at the University of Washington Information School. She is also an adjunct assistant professor in the Department of Sociology, an affiliate of the UW Center for Statistics and the Social Sciences. She studies online communication and information-related behaviors in the context of emergencies and disaster events. Her work also explores the structure and dynamics of interpersonal and organizational networks in both online and offline environments.

Zack W. Almquist is an assistant professor of sociology and statistics at the University of Minnesota. He is affiliated with the Minnesota Population Center, The Life Course Center, and the Data Science program. His interests include social network analysis, big data, computational social sciences, mathematical sociology, methodology, spatial analysis, demography, public health, and human judgment and decision making. His main research focus is on understanding, modeling, and predicting the effects space (geography) and time have on human interaction (e.g., friendship or needle sharing) and outcomes (e.g., disease or knowledge).

Carter T. Butts is currently a professor in the Departments of Sociology, Statistics, and EECS and the Institute for Mathematical Behavioral Sciences at the University of California, Irvine. His research involves the development and application of mathematical, computational, and statistical techniques to theoretical and methodological problems within the areas of social and biological network analysis, mathematical sociology, quantitative methodology, and human judgment and decision making. Currently, his work focuses on the structure of spatially embedded large-scale interpersonal networks; theory and applications of discrete exponential family models for cross-sectional and dynamic relational data; Bayesian methods for network inference, model integration, and properties of biological systems; relational event models; network sampling; and methods for the analysis of time-resolved activity data.