4-21-2014

Tie Strength, Embeddedness, and Social Influence: A Large-Scale Networked Experiment

Sinan Aral
New York University

Dylan Walker
Chapman University, dylan@chapman.edu

Tie Strength, Embeddedness & Social Influence: Evidence from a Large Scale Networked Experiment

Sinan Aral
NYU Stern School of Business & MIT, 44 West 4th Street, Room 8-81, New York, NY 10012.
sinan@stern.nyu.edu

Dylan Walker
NYU Stern School of Business, 44 West 4th Street Room: 8-80, New York, NY 10012
dwalker@stern.nyu.edu

Understanding peer influence in networks is critical to estimating product demand and diffusion, creating effective viral marketing, and designing 'network interventions' to promote positive social change. But several statistical challenges make it difficult to econometrically identify peer influence in networks. Though some recent studies use experiments to identify influence, they have not investigated the social or structural conditions under which influence is strongest. We investigate the two most prominent network characteristics that may moderate social influence between peers -- tie strength and network embeddedness. By randomly manipulating messages sent by adopters of a Facebook application to their 1.3 Million peers, we were able to identify the moderating effect of tie strength and embeddedness on influence. We find that both embeddedness and tie strength increase influence. Individuals experience a 0.6% increase in influence over their peers for each friend they share in common with that peer. As the number of common friends can be quite large, this effect is also economically significant. Individuals exert 125% more influence on peers for each affiliation they share in common, 1355% more influence on peers with whom they attended the same college, and 622% more influence on peers that live in the same current town. However, the amount of physical interaction between friends, measured by co-appearance in photos, does not have an effect. This work presents some of the first large scale experimental evidence investigating the social and structural moderators of peer influence in networks. The results could enable more effective marketing strategies and public policy more broadly.

Key words: Peer Influence, Social Contagion, Social Networks, Viral Marketing, Information Systems, Randomized Experiment.

---

1. Introduction

Social influence in networks is recognized as a key factor in the propagation of ideas, behaviors, and economic outcomes in society. Understanding the role social influence plays in spreading economic behaviors is critical to constructing sound policy in both the public and private sectors. Fortunately, emerging online systems that increasingly connect people and mediate their interactions also provide opportunities to acquire micro-level data at population scale (e.g., Eagle, M. Macy, and Claxton 2010; Golder and M. W. Macy 2011) and to conduct experiments that address endogeneity (e.g., Aral and Walker 2011a; Bakshy et al. 2012). These two advantages can be leveraged to create new experimental analytical methods that yield real-time, context-specific inferences about the role of social influence in consumer demand, marketing and public policy. Moreover, large-scale data from randomized experiments permits the detection of nuanced or subtle effects that are economically important but difficult to observe with observational analytics. It is precisely these nuanced effects that are in danger of being eclipsed by bias in endogenous processes, making randomized experimentation in large-scale systems a vital tool in the arsenal of modern business analytics.

A primary question in understanding the role of social influence in the diffusion of new products, ideas, behaviors and outcomes is how heterogeneity in the relationships between individuals impacts the level of influence they exert on one another. Despite decades of observational research, results in this domain remain inconsistent and elusive. This is perhaps unsurprising given notoriously difficult statistical challenges like simultaneity (Godes and Mayzlin 2004), homophily (Aral, Muchnik, and Sundararajan 2009), unobserved heterogeneity (Van den Bulte and G. L. Lilien 2001), time-varying factors (Van den Bulte and G. L. Lilien 2001), and other contextual and correlated effects (Manski 1993) that make it difficult to distinguish causal peer influence from other confounds that lead to behavioral clustering in network space and time. Recent research has employed exogenous shock methods to identify influence in the presence of these confounding factors (Tucker 2008). Some new methods separate peer influence from homophily and other confounding factors in observational data (Aral et al. 2009), which is useful because most data on these questions is observational. But controlling for unobservable factors, such as latent homophily, remains difficult (Shalizi and Thomas 2011).

As an alternative to observational analysis, experimental network studies using random assignment can provide a more robust means of identifying causal peer effects in networks and distinguishing influence from confounding factors. Some recent experiments have demonstrated a role for peer influence in product adoption (Aral and Walker 2011a, 2012; Bakshy et al. 2012; Bapna and Umyarov 2012), health behaviors (Centola 2011) and altruism (Leider et al. 2009). Though these studies use experiments to address confounds and identify peer influence in different network contexts, they have not investigated the social or structural conditions under which influence is strongest, an area identified as a critical new frontier in the science of social influence (Aral 2012). Two of the most widely studied factors that are likely to affect the strength of social influence are *embeddedness*, the extent to which individuals share common peers, and *tie strength*, the significance or intensity of the relationships between individuals. We investigate how embeddedness and tie strength moderate social influence in product adoption, while simultaneously controlling for confounding factors that can bias inference in networked settings .

We conducted a randomized trial of social influence in the adoption of a commercial application amongst 1.3M users of the popular online social network Facebook.com. Using novel techniques of randomized experimentation in networked environments and statistical analysis, we simultaneously identify and distinguish influence-driven outcomes from spontaneous outcomes and examine the role that social embeddedness and tie strength play in the level of influence exerted between individuals and their peers. We extend the definition of *tie strength* and examine several well-defined measures of the strength of ties (SoT) that describe the nature of the relationship between individuals and their peers in a concrete manner, including a) the social context of the relationship - how individuals met, know one another or interact with each other (e.g. whether peers attended the same college, come from the same hometown, or share common institutional affiliations), b) the recency of the relationship (e.g. whether peers currently live in the same town), c) the overlap of common interests (e.g. being fans of the same Facebook pages, joining the same Facebook groups), and d) frequency of the interaction (e.g. co-presence in photos online). This work builds upon prior research on the role that individual attributes play in social influence processes (Aral and Walker 2012). As prior research demonstrated that, in spreading processes, not all individuals are created equal; our work demonstrates that not all relationships are created equal. Our approach can be generalized and extended with relative ease to a multitude of systems of interest to researchers across the fields of marketing, management, information systems and other quantitative social sciences, where large-scale randomized field experiments are rapidly gaining traction as a powerful analytical tool.

2. Theory

2.1 Social Influence

Understanding how word-of-mouth (WOM) "buzz" about a product, service, opinion or behavior can impact its adoption has long been considered crucial to how firms can promote diffusion (Arndt 1967; Brown and Reingen 1987; Engel, Kegerreis, and Blackwell 1969; Godes and Mayzlin 2004; Katz and Lazarsfeld 1955; Manchanda, Xie, and Youn 2008). Traditionally, WOM has been specified as information (often about opinions, preferences or choices) deliberately exchanged through face-to-face interactions, though more recently the term has been applied to online or technology-enabled information exchange between individuals or from one individual to a group of others (as in the case of consumer product reviews). In contrast, researchers in economics and marketing have employed the phrase "observational learning" to refer to circumstances where a consumer observes (usually) aggregated decisions of a population prior to making his selection from a set of alternatives (Banerjee 1992; Bikhchandani, Hirshleifer, and Welch 1998; Zhang 2010).

Yet the distinction between WOM and observational learning is not always clear. For example, many online social networks automatically disseminate information about an individual's actions to his immediate peers (e.g. "Brian is reading *A Tale of Two Cities*") in a way that does not indicate aggregate decisions in larger populations or preferences amongst a set of clear alternatives. Social exchanges such as these straddle the boundary between WOM and observational learning. Notions of customer intent, opinion, preference and content valence that are typically studied in WOM research may be ambiguous or subjective in this type of information exchange and observational learning models may not be well suited to situations in which the set of alternatives is very large and information about peer decisions and outcomes are sparsely distributed.

As Godes et al. recently highlighted, the conventional definition of WOM is not suitable for a variety of social interactions that mediate information and influence. They instead proposed the more encompassing term *social influence* to describe "an action or actions taken by an individual not actively engaged in selling the product or service that impacts others' expected utility for that product or service" (Godes et al. 2005). In the remainder of this section, we describe and expand on the specific notions of the *channel, content* and *impact* of social influence proposed by Godes et al. In a later section, we discuss how these theoretical dimensions of social influence informed and guided the design of our randomized field experiment.

The *channel of social influence* refers to the medium through which influence is communicated or transmitted. Several dimensions specify the channel, such as the number of senders and recipients involved, which may be one-to-one (as in the case of personal email), one-to-many (as in the cases of email lists, online recommendations, group invitations, and automated peer referrals), many-to-many (as in the case of polls in online community forums), or many-to-one (as in the case of voting on online forum comments). Other salient dimensions include how the recipients are selected, the credibility of the channel, and whether the channel is mediated by a third party. The *content of social influence* refers to the information that is transmitted over the channel. For example, information can include individual decisions or outcomes relating to product features or product adoption, factual information about product features, or subjective opinions about the product as in the case of peer recommendations or customer reviews. Salient dimensions of the content are the subjectivity (fact vs. opinion) and whether the content is personalized to the intended recipients. Finally, the *impact of social influence* refers to the overall effect social influence may have on the actions of others. Salient dimensions of impact primarily relate to how it is measured, e.g. whether the impact is inferred or measured directly and what it means to "impact" an outcome. From our perspective, a key dimension of impact is the causal effect of an individual on their peers' behavior. As Aral (2011: 217) has argued, defining social influence as creating behavior change or "[h]ow the behaviors of one's peers change the likelihood that (or extent to which) one engages in a behavior," is essential to making effective marketing and public policy decisions because effective policy requires an understanding of how behavior is likely to change as a result of an intervention.

The specification of social influence we outline is useful in relating existing research on differing forms of social influence to one another. Moreover, it informs the design of our experiment. We use firm-mediation to control the delivery of automated notifications with impersonal content to randomly selected peer targets and assess the impact of social influence by directly measuring peer adoption response. We discuss the ramifications of these design choices in more detail in the experimental design section.

2.2 Influence and Susceptibility

Prior research in social influence has primarily focused on the dual notions of influence and susceptibility (or influenceability). The idea that some individuals are more influential than others and therefore play a catalyzing role in spreading opinions, innovations and products (Van den Bulte and Joshi 2007; J. Coleman, Katz, and Menzel 1957; Gladwell 2002; Rogers 2003; T. W. Valente 1995) is sometimes referred to as "the influentials hypothesis." Other research, focusing on the complementary idea that individual *susceptibility to influence* is the dominant driving mechanism behind diffusion in social networks, is represented in a variety of theoretical threshold-based contagion models in which behavior adoption occurs when some number or proportion of one's peers have adopted beyond one's intrinsic adoption threshold (e.g., Granovetter 1978; T. Valente 1996; Watts and Dodds 2007).[1] Though studies estimating the importance of influentials and susceptibles in the diffusion of products or behaviors in real-world networks significantly lag theoretical and simulation models of influence-based contagion, a recent observational study examined the combined notions of influence and susceptibility in social influence processes (Iyengar, Van den Bulte, and T. W. Valente 2011). More recent work has empirically identified individual influence and susceptibility, demonstrated that both mechanisms together determine the propa-

[1] In this context, a susceptible individual is one with a low intrinsic threshold.

gation of behaviors in social networks and also explored *dyadic influence*, in which the influence exerted by an individual on their peer depends on dyadic or pairwise characteristics of both parties (Aral and Walker 2012). We unify these theoretical constructs and empirically measure dyadic influence arising from heterogeneity in embeddedness and tie strength while controlling for heterogeneity in individual influence and susceptibility as well as for tendencies toward non-influenced spontaneous adoption.

2.2 Impact of Social Embeddedness and Tie Strength on Social Influence

2.2.1. Embeddedness

Network embeddedness, or the number of friends that two individuals in a relationship share in common (Easley and Kleinberg 2010: 55), has long been theorized to affect the level of trust, altruism, cooperation and communication in relationships. Embedded relationships are likely to conduct greater social influence because the presence of third party ties increases the level of trust between embedded peers (Uzzi 1997). As the relationship is "on display" in a social sense, recommendations from embedded peers are likely to be truthful revelations of product experiences or the perceived benefits of the recommended product for the party receiving the recommendation (Granovetter 1985; Uzzi 1996). Embeddedness also engenders greater cooperation because news of non-cooperative behavior spreads quickly in the network making it harder for the uncooperative actor to maintain friendly ties with third parties. In this way, embeddedness enables the development of cooperative norms that facilitate mutual helping relationships (J. S. Coleman 1988; Granovetter 1985). Embedded relationships also typically create opportunities for greater knowledge transfer between individuals (Reagans and McEvily 2003)and more fine-grained information flows (Uzzi 1997) that are multifaceted in that they provide information across multiple topics or dimensions of topics (Aral and Van Alstyne 2011). For example, when discussing a product, two consumers in an embedded relationship may share more information about the product, more knowledge about different dimensions or features of the product and more fine grained knowledge of the product, its uses and its strengths and weaknesses compared to other similar products. Greater trust, cooperation and fine grained information exchange is likely to increase the influence conducted in a relationship and thus we expect embedded relationships to convey greater influence. We adopt the conventional network structural measure of embeddedness, defined in this context as the *number of common friends* shared by individuals and their peers.

2.2.2. Tie Strength

Granovetter (1973: 1361) defines tie strength as "a (probably linear) combination of the amount of time, the emotional intensity, the intimacy (mutual confiding), and the reciprocal services which characterize the tie." He notes that strong ties also typically display multiplexity or the exchange of multiple topical contents through the relationship. Strong tie relationships are more likely to conduct influence because they convey greater trust, more fine-grained information exchange and cooperation (J. S. Coleman 1988). Although tie strength is a multi-dimensional theoretical construct, most studies use measures of the frequency of interaction to proxy for the strength of ties. For example, recent work relating social influence to political mobilization concluded that strong ties were associated with greater social influence, but defined tie strength purely in terms of the frequency of online interactions between peers (Bond et al. 2012). Prior theory has defined tie strength in terms of a number of aggregate measures that include frequency of interaction, surveyed relationship category (such as "friend", "neighbor", "relative" or "acquantaince") and perceived importance or intimacy (Brown and Reingen 1987; Frenzen and Davis 1990; Granovetter 1983). These extensions of the definition of tie strength beyond interaction frequency are clearly meaningful. At the same time, survey instruments that codify the nature of relationships between individuals and their peers are subject to perception bias and such approaches do not scale to large systems. Aggregating the nature of relationship categories into a single measure is also, in some sense, undesirable as it obscures meaningful differences in the type and quality of relationships and reduces our ability to detect the impact of the different dimensions of tie strength on influence.

We expand this conceptualization of tie strength to capture several different dimensions of relationships that may be relevant to the strength of social influence. We treat tie strength as a collection of well-defined measures about the historical and current nature of the relationships between individuals and their peers and simultaneously examine the distinct impact of these different aspects of tie strength. We define the strength of ties as including a) the social context of the relationship - how individuals met, know one another or interact with each other (e.g. whether two Facebook friends attended the *same college*, come from the *same hometown*, or the *number of common institutional affiliations* they share), b) the recency of the relationship (e.g. whether two Facebook friends currently live in the *same town*), c) the overlap of common interests (e.g. the *number of common Facebook pages they are "fans"* of or the *number of common Facebook groups* they have joined), and d) frequency of the interaction (e.g. friends' *co-presence in photos* online).

While we expect greater social affiliation and interaction is predictive of greater influence conducted between friends while similarity in preferences and interests is predictive of correlations in non-influenced spontaneous adoption, we note that our analysis is exploratory rather than confirmatory. Although there are many theoretical reasons to believe that tie strength is predictive of influence, theoretical distinctions between different types of tie strength are not yet well developed enough to provide clear theoretical predictions about how they would moderate the degree of influence conducted in a given relationship. The tie strength distinctions we observe in our data therefore provide an initial exploration of the dimensions of tie strength that may relevant to influence.

2.2.3. The Endogeneity of Social Structure and Influence

The review of relevant theories of tie strength and embeddedness in the previous two sections immediately calls attention to the endogenous processes that govern them. It is therefore perhaps unsurprising that previous empirical research on embeddedness, tie strength and social influence has been hampered by endogeneity and spurious correlation. In real-world networks, social embeddedness and tie strength are often correlated with each other and with homophily, making their measurement difficult to untangle in practice (Rogers 2003). Individuals tend to form closer relationships with similar peers, close friends are more likely to share more friends in common, and friendships become stronger with shared common experiences and friends. Nonetheless tie strength, embeddedness and homophily can be clearly distinguished theoretically. While it may be less common statistically speaking, an individual and her peer may share many mutual friends, despite being dissimilar on demographic or personality dimensions. Similarly, some close friends may have non-intersecting peer groups. Endogeneity amongst these tie characteristics is exacerbated in observational and survey-based studies that do not properly control for selection bias.

Natural social influence processes often involve endogenous communication patterns.[2] Individuals select into sending, receiving or soliciting influence-mediating communications to or from their peers. As a consequence, studies of social influence, embeddedness and tie strength that do not explicitly control for selection biases in communication patterns confound our understanding of how tie characteristics moderate influence. For example, some studies on social influence and tie characteristics report that dissimilarity between peers is correlated with increased influence (e.g. Gilly et al. 1998). In contrast, other studies contend that more homophilous ties, though more likely to be activated, are not associated with any more (or less) influence (Brown and Reingen 1987). Studies have also examined the role of embeddedness in social influence. For example, recent work on influence in the decision to join Facebook in response to an email invitation indicates that less embedded ties are associated with greater influence, leading to a higher probability of positive response (Ugander et al. 2012). None of these studies account or control for selective communication. Burt highlights another potential confounding factor: weak ties are more likely to transmit novel information, by virtue of being less socially embedded (Burt 2005).[3] These considerations highlight that inferring the impact of tie strength and social embeddedness on influence is difficult because influence-mediating communications are inherently endogenous. One notable exception is provided by the work of De Bruyn and Lilien who disentangled selection bias in communication tendencies by explicitly modeling multiple stages of interaction in influence processes in an experimental study of word-of-mouth viral marketing (De Bruyn and G. Lilien 2008). However, this approach relies on the ability of researchers to correctly model the stages of interaction in social influence processes

[2] Communication patterns may include information seeking behavior (such as solicitation of peer advice), passive subscriptions to information sources (such as blogs, twitter feeds or Facebook newsfeeds), active forwarding of information to peers (such as personalized referrals or invitations).

[3] In the framework of Godes, this corresponds to selection bias in both the channel and content of social influence.

and does not generalize well to contexts in which we lack intuition about what social processes are at work.

Our study design disentangles both the impact of the frequency of interaction and the novelty of information exchanged from other aspects of tie strength that characterize the nature of the relationship between individuals and their peers by controlling the channel of influence (through randomized recipient selection) and holding message content constant. In our design, influence-mediating communication is controlled by a third party, allowing us to randomize message target selection and to homogenize content. Messages sent from individuals to their peers contain approximately the same information, allowing us to study the impact of embeddedness and tie strength holding information diversity or novelty constant. Message target randomization also ensures that the number (frequency) of influencing mediating messages sent from individuals to their peers is independent of the embeddedness and strength of ties.

Communication patterns between individuals and their peers certainly play a role in influence and in part comprise what makes individuals influential on their peers. However, understanding the impact of social embeddedness and tie strength on influence, holding communication patterns constant is important for two reasons. First, it contributes to our understanding of how recipients of an influence-mediating message would respond differently to more or less embedded peers and peers with whom they are strongly or weakly connected. Such insights are critical to viral marketing initiatives designed to target advertisements at those most likely to maximize the diffusion products and services in a population through their natural or intrinsic influence (Aral et al. 2009). Second, it can inform policies that operate outside of the scope of natural influence (such as individual and network-based interventions and peer-oriented incentive schemes), which are deliberately designed to impact and alter natural communication and information flow patterns.

3. Empirical Methods

We partnered with a firm that develops commercial applications hosted on the popular social networking website Facebook.com. A commercial Facebook application was designed and publically released in concert with the launch of the experiment. This application provides users the opportunity to share information and opinions about movies, actors, directors and the film industry in general. As adopters used the application, automated notifications were delivered to randomly selected peers in their local social networks. Data on individual attributes of adopters and their local peers, time-stamped delivery of automated notifications and subsequent time-stamped adoption responses were collected throughout the course of the experiment.

The experimental design employs *message target randomization* to deliver automated notifications to randomly selected peers of existing application users. In this scheme, packets of notifications are generated when application users take one of several actions within the Facebook application (e.g. when the user rates a movie or friends a celebrity). These notifications are then delivered to randomly chosen subsets of the application user's Facebook friends. The random selection of a set of recipient peers is performed on a per-packet basis (i.e., a different set of recipient peers is randomly chose each time a packet is sent from an application user). This design is illustrated in Figure 1, which displays a diagram of the delivery of two packets over sequential time periods.

*** Insert Figure 1 about here ***

At time *t1*, the application user performs a packet-generating action within the application and a packet of notifications is generated. At time *t2* randomly chosen peers of the application user are designated as recipients and receive the notifications. At time *t3*, the application user performs a second packet-generating action within the application and a second packet of notifications is generated and delivered at time *t4* to a (different) set of randomly chosen peer recipients. At any given time throughout the course of the experiment, peers of an application user received 0, 1, 2 or more influence-mediating messages from their application user friend. The exposure of a peer to influence-mediating messages (notifications) over time is exogenously determined as a function of the randomization procedure. Over time, peers are assigned to risk groups (corresponding to the number of influence-mediating messages received), where risk is monotonically and randomly increasing over the course of the experiment. We discuss the implication of message target randomization on our modeling strategy and censoring procedures in the section that follows.

Automated notifications (passive viral messages) have several advantages over alternate types of influence-mediating communication for the purpose of our experimental design (Aral and Walker 2011a, 2011b, 2012). First, the automated notifications channel is mediated by a third party, allowing the desired level of experimental control and randomization of peer targets. Since the recipients of notifications and the decisions of whether and when to send them are all automated, selection effects on the part of the sender that might otherwise introduce bias can be avoided.

Second, the content of automated notifications employed in the experiment included only impersonal information about the sender's use of the application (as is typical of information exchange in observational learning scenarios). The inclusion of only impersonal information in influence-mediating messages allows for the measurement of the impact of social influence while holding constant the potentially large degree of heterogeneity present in personalized, sender-created content. Heterogeneity in the content of messages created by individuals is known to have a significant effect on social influence. In particular, the effects of content heterogeneity on influence have been studied in the context of the positive valence of the message content (Berger and Milkman 2009) and the effectiveness of viral features that allow message tailoring or personalization (Aral and Walker 2011). While content heterogeneity can and in all likelihood does play a major role in what makes individuals influential, simultaneous variation of content and relationship attributes can confound measurements of the effect of relationship attributes on social influence.

Third, the delivery of notifications to only a random subset of an individual's peers permits direct comparison of the response of treated peers to peers of the same application user that were not treated. When the targets of potentially influential communications are randomized amongst peers of the same application user, any homophilous structure between an application user and his treated and untreated peers and the propensity to select a particular peer to notify are held constant and are identical for recipient and non-recipient peer groups. Other unobserved factors that could potentially drive influenced adoption, such as offline or alternative online communications, can also be cleanly distinguished with this design, because recipient and non-recipient peers in expectation share similar propensities to receive and be affected by such communications on average. Moreover, homophily in unobserved attributes (latent homophily) that may be indicated by the very existence of a relationship of peers with a common friend (see Shalizi and Thomas 2011) will be equally represented in recipient and non-recipient peer groups. Differences in adoption outcomes between recipient and non-recipient peers can then be attributed solely to the influence-mediating messages they received.

4. Analysis and Results

4.1. Data and Descriptive Statistics

Throughout the 44-day experimental period, we collected individual level profile data from 7,730 application users and their 1.3M distinct peers as well as time-stamped click stream data on notification delivery and subsequent peer responses. During this time, 41,686 automated notifications were delivered to randomly chosen peer targets of application users, resulting in 967 peer adoptions, a 13% increase in product adoption.[4] Collected user data included the social network of adopters and all mutual ties between their peers[5] and individual level profile data included *age*, *gender*, *relationship status*, *hometown*, *current town*, *college attendance*, *affiliations*, *Facebook pages*, *Facebook group membership*, and *tagged appearance in photos*[6].

4.2. Model Specification

Following prior work on influence identification and social contagion, we adopt a hazard modeling approach (Aral and Walker 2011a, 2012; Van den Bulte and G. L. Lilien 2001; Iyengar et al. 2011; Nam, Manchanda, and Chintagunta 2010). We employ a Cox proportional hazard models to estimate the hazard of a peer (of an existing application user) to adopt. The model simultaneously estimates the impact

[4] This represents an adoption rate that outperforms click thorough rates for traditional banner advertising and that is on par with click through rates for email campaigns (see Aral and Walker 2011a: 25).
[5] In addition, a global sample of the social network for 12M users was collected by combining ego-sampling and mutual ties between peers for the union of all users of Facebook applications developed by the collaborating firm.
[6] Facebook affiliations typically indicate some past experience such as working for a company or belonging to the same institution or society. Facebook pages typically indicate interest in activities, brands, products, bands, and media personalities. Facebook groups are online version of social groups that allow users to interact with one another in centralized discussion forums and view news and events pertaining to that group.

of social embeddedness and multiple measures of tie strength on influenced and spontaneous adoption while controlling for the moderating effect of individual attributes on influenced and spontaneous adoption:

$$\lambda_j(i,j,t) = \lambda_0(t)\exp\{\beta_N N_j(t) + \beta^i_{Spont}X_i + \beta^j_{Spont}X_j + \beta^{SoT}_{Pref}SoT_{ij} + \beta^{Embed}_{Pref}Embed_{ij} + \beta^i_{Infl}X_iN_j(t) + \beta^j_{Susc}X_jN_j(t) + \beta^{SoT}_{Infl}N_j(t)SoT_{ij} + \beta^{Embed}_{Infl}N_j(t)Embed_{ij}\}$$

where λ_j is the hazard for a peer *j* of a user *i* to adopt; $N_j(t)$ is the number of notifications received by a peer *j*; $Embed_{ij}$ is the embeddedness of the relationship between user *i* and peer *j* (the number of friends shared by individual *i* and peer *j*); and SoT_{ij} is a vector of the tie strength attributes characterizing the relationship between individual *i* and peer *j*. These models include a rich set of covariate controls for demographic and individual-level characteristics of individuals (X_i) and their peers (X_j) (including *age*, *gender* and *relationship status*). β_N captures the raw impact of influence holding constant heterogeneity in individual or tie attributes—it represents the marginal impact of a peer receiving influence-mediating messages irrespective of the individual attributes of individual *i* and peer *j* or their dyadic tie attributes. β^i_{Spont} captures the tendency for peers of application users with attributes X_i to spontaneously adopt in the absence of influence ($N_j = 0$). β^j_{Spont} captures the tendency for peers with own attributes X_j to spontaneously adopt in the absence of influence ($N_j = 0$). β^{SoT}_{Pref} captures the extent to which the measure of tie strength indicates a similarity in preference for a peer to adopt the product spontaneously in the absence of influence ($N_j = 0$) given that her application user friend has adopted. Similarly, β^{Embed}_{Pref} captures the extent to which peers with embedded ties to existing application users will have a preference to adopt the product spontaneously in the absence of influence ($N_j = 0$). β^i_{Infl} represents individual influence – it captures the impact of individual attributes of an application user (X_i) on the hazard of her peer to adopt due to influence per influence-mediating message received. β^j_{Susc} represents individual susceptibility to influence – it captures the impact of peer attributes (X_j) on peers' hazard to adopt due to influence per influence-mediating message received. β^{SoT}_{Infl} captures the impact of the tie strength measure (for the tie between application user *i* and peer *j*) on influence-driven adoption per influence-mediating message received. β_{Infl}^{Embed} captures the impact of the embeddedness of the tie between application user *i* and peer *j* on influence-driven adoption by *j* per influence-mediating message received. The model estimation provides good concordance with observed data and Wald, logrank, and likehood ratio test statistics indicate strong likelihood, significance, and goodness of fit (see table A1 in the appendix).

As peers of application users randomly receive (multiple) notifications from their application user friend throughout the course of the experiment, we employed interval censoring to transition users from one risk group (e.g., the hazard associated with receiving one influence-mediating notification) to the next (e.g., the hazard associated with receiving two influence-mediating notifications). Peer adoption outcomes may be correlated because peers of a given application user share a common application user friend. To account for this, the hazard model employs robust errors clustered on the identity of the application user friend.

4.3. Results

Our model specification enables us to estimate two quantities of theoretical interest: influence-based adoption and spontaneous adoption. Influence-based adoption measures the degree to which a particular relationship characteristic moderates the influence an individual has over their peers or the degree to which that characteristic is associated with changing someone's behavior from not adopting to adopting the application. Spontaneous adoption on the other hand measures the degree to which a particular relationship characteristic predicts a correlated latent preference to adopt the product. For example, if having attended the same college is associated with spontaneous adoption but not influence-based adoption, then attending the same college is predictive of preference similarities that predict adoption by a peer of a current adopter. If on the other hand having attended the same college is associated with influence-based adoption but not spontaneous adoption, then individuals influence their friends who attended the same college as they did more than their friends who attended different colleges.

These distinctions are critical to marketing policy and other general networked interventions. Predictors of spontaneous adoption can generate good sets of advertising targets whose preference similarities to current adopters make them likely to respond positively to advertisements about the product under consideration. On the other hand, moderators of influence can highlight good sets of relationship pairs in which incentives to propagate social influence may work well to create additional product adoptions (Aral and Taylor 2011). For example, incentives to 'invite friends' to products or discounts for friends and family adoption may work well in relationship pairs that conduct influence. Model estimations of the impact of relationship characteristics on influence-based and spontaneous adoption are tabulated in table 1 (for full model estimations, see table A1 in the appendix).

4.3.1. Effects of Social Embeddedness and Tie Strength on Influence

Results displaying the impact of social embeddedness and tie strength on influence are displayed in Figure 2. The forest plot displays the hazard ratios, standard errors (boxes) and 95% confidence intervals (whiskers) of influence-driven adoption associated with embeddedness (number of common friends) and tie strength attributes. The hazard ratios displayed are relative to the baseline case (a hypothetical blank social network profile) and all categorical dummy variables such as *Same college* and *Diff colleg*e catalogue all categories for which the measure is defined. The holdout sets are cases for which either the application user or peer have not reported the measure in their social network profile. For example, the three exhaustive categories for college affiliation relations are 1) the peers went to the same college, 2) the peers went to different colleges, or 3) one or both peers do not report their college attendance.

*** Insert Figure 2 about here ***

We observe several interesting patterns in the results detailing the impact of embeddedness and different measures of tie strength on influence. First, tie measures that capture peers' joint participation in common social or institutional contexts between individuals and their Facebook friends are associated with greater influence. Individuals exert 125% more influence on friends for each institutional affiliation they share in common ($p < 0.05$). Attending the same college as one's friend is associated with a 1355%

increase in influence ($p < 0.01$) compared to attending different colleges. This represents the largest impact on influence of the categorical measures of tie strength we considered. In contrast, coming from the same hometown is not significantly associated with influence, perhaps suggesting that this measure accurately captures more causal or even incidental social contexts (e.g., weak ties resulting from Facebook users' desires to keep in contact with casual acquaintances).

Second, tie strength measures associated with current or recent social contexts exhibit differing impacts on influence. Individuals exhibit 622% more influence on friends that live in the same current town ($p < 0.01$). This is interesting because ties between friends currently living in the same town may indicate joint involvement in more recent social contexts (e.g., friendships that are more recent or recently relevant). Interestingly, appearing in photos with peers, an indicator of offline interaction at significant events, is not significantly associated with influence.

Third, tie strength measures associated with common interests or preferences do not moderate influence. Individuals are no more or less influential on peers with whom they share common Facebook pages or are co-members of online groups.

Finally, individuals are more influential on peers with whom they are more embedded, exhibiting a 0.6% increase in influence for each friend they share in common ($p < 0.001$). The impact of embeddedness on influence, though comparatively smaller than the tie strength measures considered here, remains economically significant, as the number of common friends can be quite large.

These results indicate the importance of social embeddedness in influence processes and the subtlety in the relationship between influence and measures of tie strength. They also highlight the importance of using disaggregated measures of tie strength.

4.3.1. Social Embeddedness and Tie Strength as Predictors of Spontaneous Adoption

Results displaying the correlation between spontaneous (preference) adoption and tie characteristics are displayed in Figure 3.

*** Insert Figure 3 about here ***

Tie characteristics (tie strength and embeddedness) associated with spontaneous adoption of the product by peers of existing adopters indicate preference similarity between peers – the extent to which the measure captures the similarity in the (latent) preference to adopt the product when a friend has already adopted. Some tie strength measures that seem to relate to common social contexts are good predictors of preference similarity, while others are not. Sharing common affiliations or attending the same college as an adopter of the product is not significantly associated with a tendency to spontaneously adopt. However, each photo that a peer shares in common with an adopter is associated with a 1.4% increase in the hazard to adopt spontaneously ($p < 0.001$). Coming from the same hometown as a peer who has adopted the product is associated with a 105% increase in the hazard to adopt spontaneously ($p < 0.05$). This indicates that hometown may be a good (latent) proxy for individual preferences for the product. One explanation for this pattern in the results could be that current friends influence us more, but that our preference-driven behaviors are more correlated with past, non-recent social contexts. In other words, we are more influenced by friends in the same current town; but our preferences are more correlated with friends from the same hometown and with friends that currently do not live in the same town.

Each additional fan page that a peer shares in common with an adopter of the product is associated with a 0.7% increase in the hazard to adopt spontaneously ($p < 0.001$). This indicates that declared preferences and interests (not directly related to the product) capture preferences for the product. Each online group that a peer participates in with an adopter of the product is associated with a 3.4% increase in hazard to spontaneously adopt ($P<0.001$). This indicates that online social activities capture latent dimensions of preference for the product.

5. Discussion and Conclusions

The availability of micro-level data at population scales has been recognized as a crucial opportunity in the advancement of modern business analytics. At the same time, micro-level experimentation in large networked environments is a new frontier for analytics which has the potential to circumvent problematic issues of causal identification and concerns of endogeneity that have hindered our understanding of the detailed social influence process involved in the propagation of behaviors and economic outcomes. Advancement of the science of social influence is vital to both marketing strategy and public policy where firms or governments seek to leverage social influence to encourage the spread of products or promote positive behaviors while curtailing negative ones. Research on social influence has predominantly focused on whether influence plays a role in the diffusion of a product or behavior and the relative size of the effect. However, recent focus has shifted to examining when and under what individual, social and structural conditions influence is stronger or weaker. This latter focus, which we adopt in this study, is important for policy as it can reveal which relationships warrant viral incentives, social interventions, targeting or network-based marketing.

We conducted a large-scale randomized experiment to identify the impact of tie strength characteristics and social embeddedness on influence. This work presents some of the first large scale experimental evidence investigating the social and structural moderators of peer influence in networks. Results from our study shed light on the role that relationship and social structural characteristics play on influence-based propagation. We found that relationship characteristics which capture joint participation in institutional or common social contexts were associated with the greatest influence. For example, individuals exerted an over thirteen-fold increase in influence over peers with whom they attended the same college. Some measures of the recency of social context in the relationships between individuals and their peers were associated with increased influence, while others were not. For example, individuals exerted an over six-fold increase in influence over peers currently living in the same town, but did not exert more influence over peers with whom they co-appeared in online photos. Interestingly, measures of tie strength based on common interests were not associated with influence, though were good predictors of preference similarity in the adoption of the product we studied. Finally, individuals exhibited greater influence on peers with whom they shared embedded relationships. This latter effect was both subtle and economically significant, highlighting the importance of large-scale randomized experiments in the detection of nuanced effects in the face of endogenity, bias and confounding factors.

Future work should address additional contexts, where the correlation between social structural measures and influence may differ. In addition, focus on alternative social measures, such as structural equivalence and brokerage may yield meaningful insights into the dynamics of social influence. The methods used in this study can be generalized to a wide variety of contexts to further our understanding of social contagions and better inform data-driven decisions in several policy domains including marketing, public health, and politics.

References

Aral, Sinan. 2012. "Social science: Poked to vote." *Nature* 489(7415):212–214.

Aral, Sinan, and Marshall W. Van Alstyne. 2011. "The Diversity-Bandwidth Tradeoff." *American Journal of Sociology* 117(1):90–171.

Aral, Sinan, Lev Muchnik, and Arun Sundararajan. 2009. "Distinguishing influence-based contagion from homophily-driven diffusion in dynamic networks." *Proceedings of the National Academy of Sciences* 106(51):21544–21549.

Aral, Sinan, and Sean Taylor. 2011. "Viral Incentive Systems: A Randomized Field Experiment." *Working Paper, Stern School of Business, NYU.*

Aral, Sinan, and Dylan Walker. 2010. "Creating Social Contagion Through Viral Product Design: A Randomized Trial of Peer Influence in Networks." *Social Science Research Network Working Paper Series.*

Aral, Sinan, and Dylan Walker. 2011a. "Creating Social Contagion Through Viral Product Design: A Randomized Trial of Peer Influence in Networks." *Management Science* 57(9):1623–1639.

Aral, Sinan, and Dylan Walker. 2011b. "Identifying Social Influence in Networks Using Randomized Experiments." *IEEE Intelligent Systems* 26(5):91–96.

Aral, Sinan, and Dylan Walker. 2012. "Identifying Influential and Susceptible Members of Social Networks." *Science* 337(6092):337–341.

Arndt, Johan. 1967. "Role of Product-Related Conversations in the Diffusion of a New Product." *Journal of Marketing Research* 4(3):291–295.

Bakshy, Eytan, Itamar Rosenn, Cameron Marlow, and Lada Adamic. 2012. "The Role of Social Networks in Information Diffusion."

Banerjee, Abhijit V. 1992. "A Simple Model of Herd Behavior." *The Quarterly Journal of Economics* 107(3):797–817.

Bapna, Ravi, and Akhmed Umyarov. 2012. "Are Paid Subscriptions on Music Social Networks Contagious? A Randomized Field Experiment." *Working paper*.

Berger, Jonah A., and Katherine L. Milkman. 2009. "Social Transmission, Emotion, and the Virality of Online Content." *Social Science Research Network Working Paper Series*.

Bikhchandani, Sushil, David Hirshleifer, and Ivo Welch. 1998. "Learning from the Behavior of Others: Conformity, Fads, and Informational Cascades." *The Journal of Economic Perspectives* 12(3):151–170.

Bond, Robert M. et al. 2012. "A 61-million-person experiment in social influence and political mobilization." *Nature* 489(7415):295–298.

Brown, Jacqueline Johnson, and Peter H. Reingen. 1987. "Social Ties and Word-of-Mouth Referral Behavior." *The Journal of Consumer Research* 14(3):350–362.

De Bruyn, A, and G Lilien. 2008. "A multi-stage model of word-of-mouth influence through viral marketing." *International Journal of Research in Marketing* 25(3):151–163.

Van den Bulte, Christophe, and Yogesh V. Joshi. 2007. "New Product Diffusion with Influentials and Imitators." *Marketing Science* 26(3):400–421.

Van den Bulte, Christophe, and Gary L. Lilien. 2001. "Medical Innovation Revisited: Social Contagion versus Marketing Effort." *The American Journal of Sociology* 106(5):1409–1435.

Burt, Ronald S. 2005. *Brokerage and Closure: An Introduction to Social Capital*. Oxford University Press, USA.

Centola, Damon. 2011. "An Experimental Study of Homophily in the Adoption of Health Behavior." *Science* 334(6060):1269–1272.

Coleman, James, Elihu Katz, and Herbert Menzel. 1957. "The Diffusion of an Innovation Among Physicians." *Sociometry* 20(4):253–270.

Coleman, James S. 1988. "Free Riders and Zealots: The Role of Social Networks." *Sociological Theory* 6(1):52–57.

Eagle, Nathan, Michael Macy, and Rob Claxton. 2010. "Network Diversity and Economic Development." *Science* 328(5981):1029–1031.

Easley, David, and Jon Kleinberg. 2010. *Networks, Crowds, and Markets*. Cambridge University Press.

Engel, James F., Robert J. Kegerreis, and Roger D. Blackwell. 1969. "Word-of-Mouth Communication by the Innovator." *The Journal of Marketing* 33(3):15–19.

Frenzen, Jonathan K., and Harry L. Davis. 1990. "Purchasing Behavior in Embedded Markets." *The Journal of Consumer Research* 17(1):1–12.

Gilly, Mary C., John L. Graham, Mary Finley Wolfinbarger, and Laura J. Yale. 1998. "A Dyadic Study of Interpersonal Information Search." *Journal of the Academy of Marketing Science* 26(2):83–100.

Gladwell, Malcolm. 2002. *The Tipping Point: How Little Things Can Make a Big Difference*. New editio. Abacus.

Godes, David et al. 2005. "The Firm's Management of Social Interactions." *Marketing Letters* 16(3/4).

Godes, David, and Dina Mayzlin. 2004. "Using Online Conversations to Study Word-of-Mouth Communication." *Marketing Science* 23(4):545–560.

Golder, Scott A., and Michael W. Macy. 2011. "Diurnal and Seasonal Mood Vary with Work, Sleep, and Daylength Across Diverse Cultures." *Science* 333(6051):1878–1881.

Granovetter, Mark. 1978. "Threshold Models of Collective Behavior." *American Journal of Sociology* 83(6):1420–1443.

Granovetter, Mark. 1983. "The Strength of Weak Ties: A Network Theory Revisited." *Sociological Theory* 1:201–233.

Granovetter, Mark. 1985. "Economic Action and Social Structure: The Problem of Embeddedness." *The American Journal of Sociology* 91(3):481–510.

Iyengar, Raghuram, Christophe Van den Bulte, and Thomas W. Valente. 2011. "Opinion Leadership and Social Contagion in New Product Diffusion." *Marketing Science*.

Katz, Elihu, and Paul Lazarsfeld. 1955. *Personal influence: The part played by people in the flow of mass communications*. Free Press.

Leider, Stephen, Markus M. Möbius, Tanya Rosenblat, and Quoc-Anh Do. 2009. "Directed Altruism and Enforced Reciprocity in Social Networks." *The Quarterly Journal of Economics* 124(4):1815–1851.

Manchanda, Puneet, Ying Xie, and Nara Youn. 2008. "The Role of Targeted Communication and Contagion in Product Adoption." *Marketing Science* 27(6):961–976.

Manski, Charles F. 1993. "Identification Problems in the Social Sciences." *Sociological Methodology* 23.

Nam, Sungjoon, Puneet Manchanda, and Pradeep K. Chintagunta. 2010. "The Effect of Signal Quality and Contiguous Word of Mouth on Customer Acquisition for a Video-on-Demand Service." *Marketing Science* 29(4):690–700.

Reagans, Ray, and Bill McEvily. 2003. "Network Structure and Knowledge Transfer: The Effects of Cohesion and Range." *Administrative Science Quarterly* 48(2):240–267.

Rogers, Everett M. 2003. *Diffusion of Innovations, 5th Edition*. 5th ed. Free Press.

Shalizi, Cosma Rohilla, and Andrew C. Thomas. 2011. "Homophily and Contagion Are Generically Confounded in Observational Social Network Studies." *Sociological Methods & Research* 40(2):211–239.

Tucker, Catherine. 2008. "Identifying Formal and Informal Influence in Technology Adoption with Network Externalities." *Management Science* 54(12):2024–2038.

Ugander, Johan, Lars Backstrom, Cameron Marlow, and Jon Kleinberg. 2012. "Structural diversity in social contagion." *Proceedings of the National Academy of Sciences* 109(16):5962–5966.

Uzzi, Brian. 1996. "The Sources and Consequences of Embeddedness for the Economic Performance of Organizations: The Network Effect." *American Sociological Review* 61(4):674–698.

Uzzi, Brian. 1997. "Social Structure and Competition in Interfirm Networks: The Paradox of Embeddedness." *Administrative Science Quarterly* 42(1):35–67.

Valente, T. 1996. "Social network thresholds in the diffusion of innovations." *Social Networks* 18(1):69–89.

Valente, Thomas W. 1995. *Network Models of the Diffusion of Innovations (Quantitative Methods in Communication Subseries)*. Hampton Press (NJ).

Watts, Duncan J., and Peter Sheridan Dodds. 2007. "Influentials, Networks, and Public Opinion Formation." *Journal of Consumer Research* 34(4):441–458.

Zhang, Juanjuan. 2010. "The Sound of Silence: Observational Learning in the U.S. Kidney Market." *Marketing Science* 29:315–335.

Tables and Figures

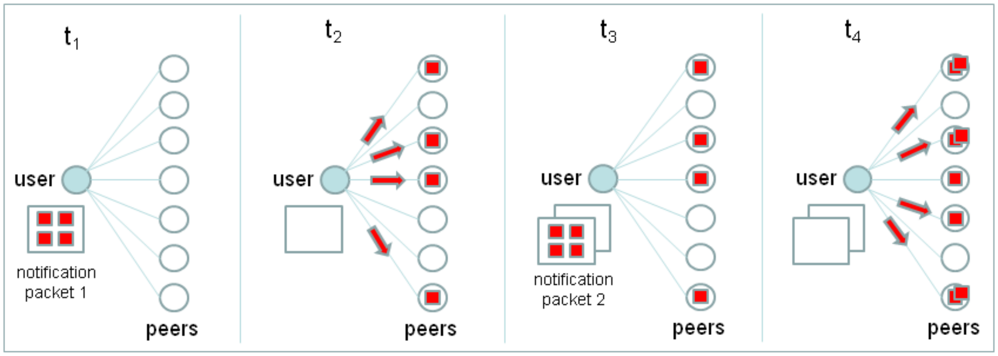
Table 1: Impact of Embeddedness and SoT on Influence

	Influence	Spontaneous
	Hazard Ratio	Hazard Ratio

	(SE)	(SE)
	β_{Infl}^{Embed}	β_{Pref}^{Embed}
Num. common friends	1.0063*** (0.0016)	1.0077*** (0.0008)
	β_{Infl}^{SoT}	β_{Pref}^{SoT}
Hometown (same)	1.0171 (0.2272)	1.4724 (0.2568)
Hometown (different)	1.3735 (0.1684)	0.7187 (0.2361)
Current town (same)	2.2899*** (0.3094)	0.4686 (0.7221)
Current town (different)	0.3171 (0.6699)	1.9300 (0.3418)
College (same)	8.5540*** (0.8389)	0.3646 (1.1272)
College (different)	0.5878 (0.4956)	0.7664 (0.3529)
Num. common affiliations	2.2548** (0.3740)	0.8184 (0.3288)
Num. common pages	1.0031 (0.0023)	1.0067*** (0.0010)
Num. common groups	1.0074 (0.0057)	1.0335*** (0.0044)
Num. photos together	0.9977 (0.0031)	1.0142*** (0.0018)

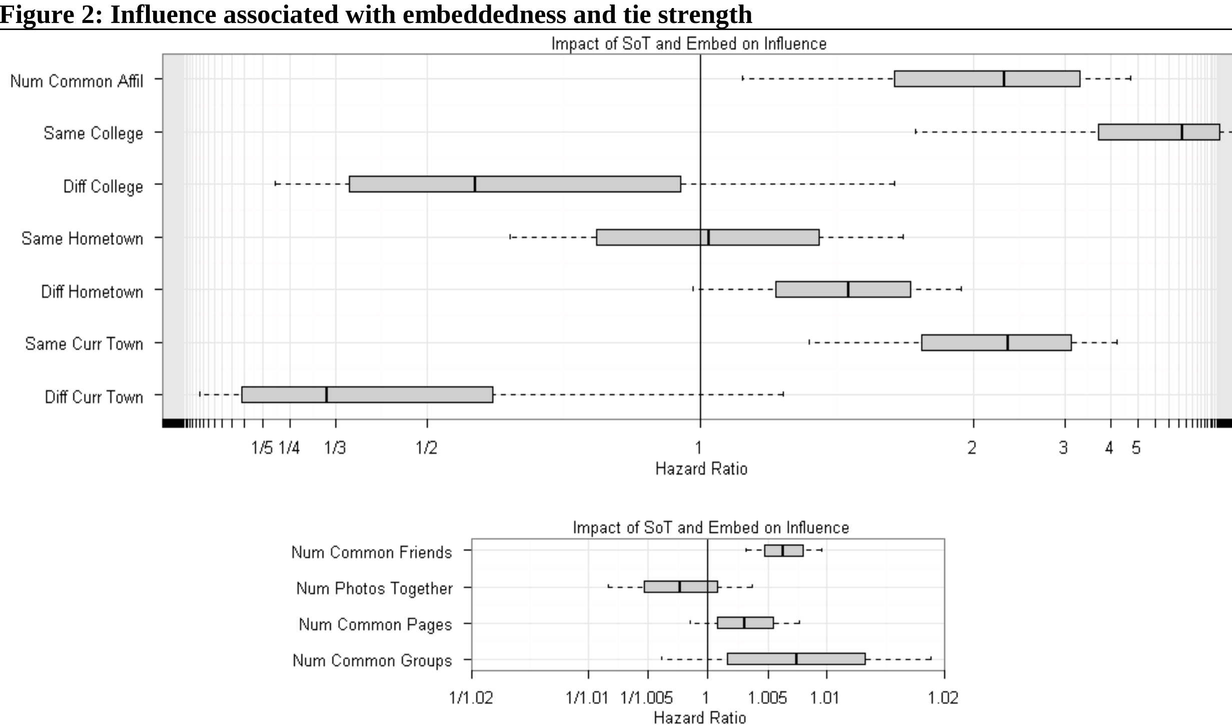
Notes: The table reports parameter estimates and standard errors from the Single Failure Proportional Hazards Model specified on page ?. Variables reported include Categorical dummy variables indicating: Hometown (same, different): whether the individual and peer come from the same or different hometowns (unreported hometown corresponds to the holdout); Current town (same, different): whether the individual and peer live in the same or different current towns (unreported current town corresponds to the holdout); College (same, different): whether the individual and peer attended the same or different colleges (unreported college corresponds to the holdout); Num. common pages: the number of Facebook pages shared in common between the individual and their peer; Num. common groups: the number of Facebook groups shared in common between the individual and their peer; Num. photos together: the number of photos in which both the individual and peer appear; Hazard ratios in the influence column correspond to variables crossed with N_j (the number of notifications received by the peer) and indicate the effect of attribute-driven adoption; Hazard ratios in the spontaneous column correspond to uncrossed variables and represent spontaneous or preference-related adoption. Statistical Significance of parameters is reported as follows: ***p<.01; **p<.05; *p<.10.

Figure 1: Randomized targeting of influence-mediating messages



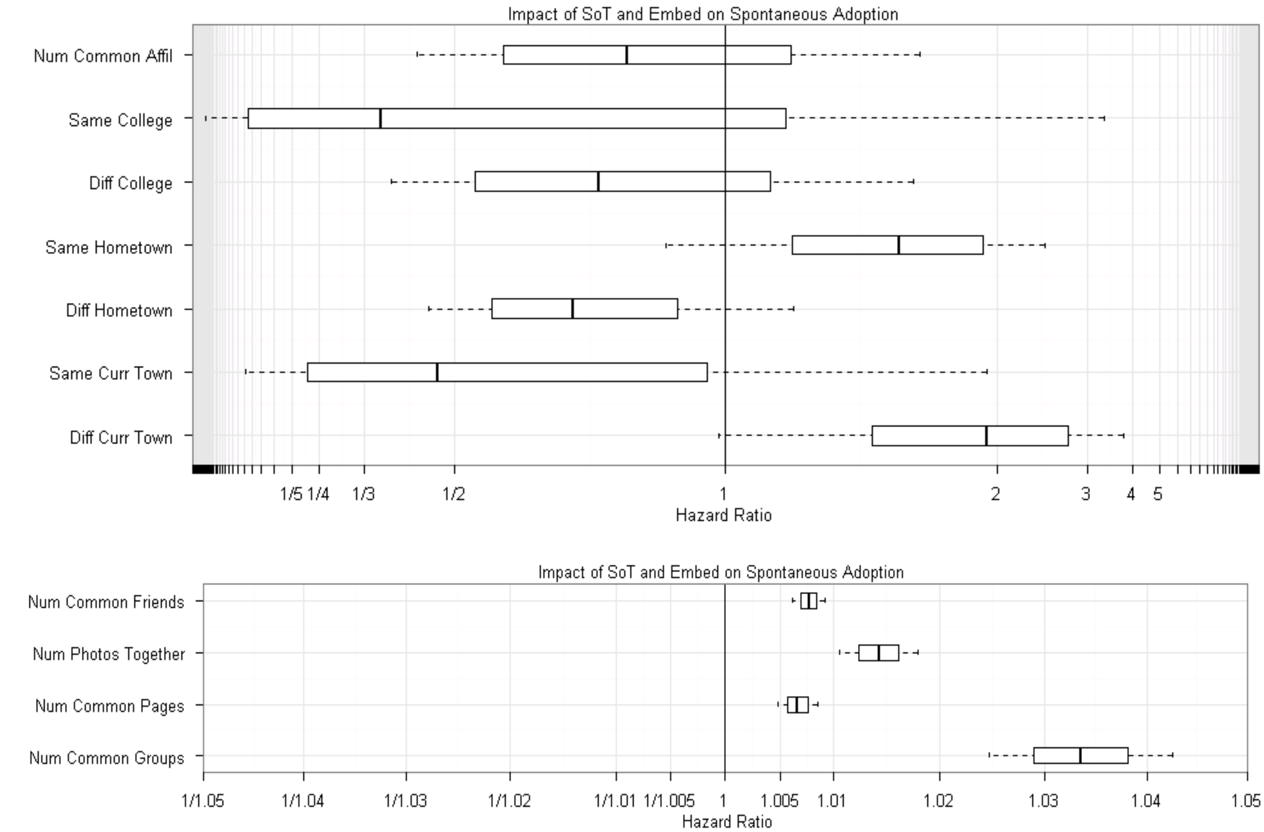
A diagram depicting the message target randomization employed in the experiment. Notification packets are generated when an application user takes a packet-generating action within the Facebook application. For each packet that is generated, the notifications in the packet are distributed to a randomly selected subset of the application user's peers. Above displays two sequential packet distributions. Different recipient targets are randomly chosen at the time of distribution for each packet.

Figure 2: Influence associated with embeddedness and tie strength



Effects of tie strength and embeddedness and tie strength measures on influenced adoption. The figure displays hazard ratios (HRs) representing the percentage increase (HR>1) or decrease (HR<1) in adoption hazards associated with each tie attribute. Boxes are standard errors. Whiskers are 95% confidence intervals.

Figure 3: Spontaneous (preference) adoption associated with embeddedness and tie strength



Effects of tie strength and embeddedness and tie strength measures on spontaneous (preference) adoption. The figure displays hazard ratios (HRs) representing the percentage increase (HR>1) or decrease (HR<1) in spontaneous adoption hazards associated with each tie attribute. Boxes are standard errors. Whiskers are 95% confidence intervals.

Appendix

Table A1: Full Estimates from the Cox Proportional Hazards Model of Influence

	β	exp (β)	$se(\beta)$	z	Pr(>|z|)	CI Lower .95	CI Upper .95
			Treatment **(β_N)**				
# Notifications	1.684494	5.389723	0.080309	20.975	< 2e-16	4.60476	6.3085
			Spontaneous Adoption of i **(β^i_{Spont})**				
Age (0-18)	0.268952	1.308592	0.18174	1.48	0.138906	0.91645	1.8685
Age (18-23)	-0.57042	0.56529	0.276804	-2.061	0.03933	0.32859	0.9725
Age (23-31)	-0.18764	0.828914	0.266578	-0.704	0.481506	0.49158	1.3977
Age (>31)	-0.021	0.979215	0.174338	-0.12	0.904104	0.69579	1.3781
Male	0.045045	1.046075	0.191168	0.236	0.813718	0.71919	1.5215
Female	0.159938	1.173438	0.150431	1.063	0.287692	0.8738	1.5758
Single	-0.22938	0.795026	0.174942	-1.311	0.189797	0.56425	1.1202
Relationship	-0.21854	0.80369	0.262786	-0.832	0.405617	0.48018	1.3452
Engaged	-0.32077	0.725594	0.450332	-0.712	0.476288	0.30017	1.7539
Married	-0.35325	0.7024	0.190385	-1.855	0.063531	0.48365	1.0201
Its Complicated	-0.09331	0.91091	0.428421	-0.218	0.827583	0.39337	2.1093
			Spontaneous Adoption of j **(β^j_{Spont})**				
Age (0-18)	-0.05414	0.947302	0.142402	-0.38	0.703815	0.7166	1.2523
Age (18-23)	0.000206	1.000206	0.152298	0.001	0.998921	0.74208	1.3481
Age (23-31)	-0.39789	0.671739	0.180174	-2.208	0.02722	0.47188	0.9562
Age (>31)	0.487589	1.628385	0.132142	3.69	0.000224	1.25683	2.1098
Male	0.392264	1.480328	0.125287	3.131	0.001743	1.15801	1.8924
Female	0.756543	2.130897	0.109676	6.898	5.27E-12	1.71872	2.6419
Single	0.163878	1.178071	0.140326	1.168	0.242873	0.8948	1.551
Relationship	-0.10779	0.897817	0.187971	-0.573	0.566349	0.62114	1.2977
Engaged	-0.39365	0.674588	0.366943	-1.073	0.283364	0.32862	1.3848
Married	0.319573	1.37654	0.167534	1.908	0.056454	0.99125	1.9116
Its Complicated	-0.58725	0.555856	0.549704	-1.068	0.285387	0.18926	1.6326
			Spontaneous (Preference) Adoption of tie i- j **($\beta^{SoT}_{Pref}, \beta^{Embed}_{Pref}$)**				
Hometown (same)	0.386861	1.472352	0.256799	1.506	0.131945	0.89007	2.4356
Hometown (different)	-0.33028	0.718724	0.236089	-1.399	0.161826	0.45248	1.1416
Current town (same)	-0.75802	0.468593	0.722129	-1.05	0.293855	0.11379	1.9296
Current town (different)	0.657508	1.929977	0.341757	1.924	0.054367	0.98775	3.771
College (same)	-1.00886	0.364633	1.127244	-0.895	0.370797	0.04003	3.3218
College (different)	-0.26611	0.766351	0.352884	-0.754	0.450781	0.38375	1.5304
Num. common affiliations	-0.20043	0.818376	0.328813	-0.61	0.542149	0.4296	1.559
Num. common pages	0.006641	1.006663	0.000951	6.986	2.83E-12	1.00479	1.0085
Num. common groups	0.032985	1.033535	0.00441	7.48	7.45E-14	1.02464	1.0425

Num. photos together	0.014124	1.014224	0.001834	7.702	1.34E-14	1.01059	1.0179
Num. common friends	0.007649	1.007678	0.000761	10.05	< 2e-16	1.00618	1.0092
Influence **(β_{Infl})**							
Age (0-18)	-0.32031	0.725928	0.149929	-2.136	0.032648	0.5411	0.9739
Age (18-23)	0.172528	1.188305	0.174764	0.987	0.323542	0.84366	1.6737
Age (23-31)	-0.06347	0.938506	0.299367	-0.212	0.832105	0.52193	1.6876
Age (>31)	0.148318	1.159882	0.198598	0.747	0.455167	0.7859	1.7118
Male	0.133022	1.142276	0.16452	0.809	0.418774	0.82743	1.5769
Female	-0.20393	0.815522	0.139504	-1.462	0.143795	0.62043	1.072
Single	0.235396	1.265409	0.216139	1.089	0.276113	0.82842	1.9329
Relationship	-0.09242	0.911721	0.262322	-0.352	0.724597	0.54522	1.5246
Engaged	-0.01265	0.98743	0.311268	-0.041	0.967583	0.53648	1.8174
Married	0.650194	1.915912	0.153836	4.227	2.37E-05	1.4172	2.5901
Its Complicated	-0.28242	0.753959	0.300552	-0.94	0.347391	0.41833	1.3589
Susceptibility **(β_{Susc})**							
Age (0-18)	0.113955	1.120702	0.098007	1.163	0.244943	0.92484	1.358
Age (18-23)	-0.20161	0.81741	0.10981	-1.836	0.066353	0.65913	1.0137
Age (23-31)	-0.0975	0.907105	0.084742	-1.151	0.249931	0.76829	1.071
Age (>31)	-0.17758	0.83729	0.087427	-2.031	0.042232	0.70544	0.9938
Male	-0.20719	0.812865	0.066524	-3.115	0.001842	0.7135	0.9261
Female	-0.35571	0.700675	0.068325	-5.206	1.93E-07	0.61286	0.8011
Single	0.313367	1.368024	0.105858	2.96	0.003074	1.1117	1.6834
Relationship	0.255853	1.291562	0.172659	1.482	0.138382	0.92076	1.8117
Engaged	0.725132	2.065003	0.221963	3.267	0.001087	1.33655	3.1905
Married	-0.00476	0.99525	0.140076	-0.034	0.972882	0.75631	1.3097
Its Complicated	0.774685	2.169909	0.294186	2.633	0.008455	1.21907	3.8624
Influenced Adoption of tie i- j **($\beta_{Infl}^{SoT}, \beta_{Infl}^{Embed}$)**							
Hometown (same)	0.016988	1.017134	0.227245	0.075	0.940407	0.65155	1.5879
Hometown (different)	0.317388	1.373535	0.168374	1.885	0.059428	0.98746	1.9106
Current town (same)	0.828517	2.28992	0.309416	2.678	0.007413	1.24866	4.1995
Current town (different)	-1.14863	0.31707	0.669876	-1.715	0.086401	0.0853	1.1786
Colle (same)	2.146397	8.55398	0.838897	2.559	0.01051	1.65233	44.2832
College (different)	-0.5313	0.58784	0.495577	-1.072	0.283681	0.22255	1.5527
Num. common affiliations	0.813058	2.254792	0.374014	2.174	0.029715	1.08329	4.6932
Num. common pages	0.003098	1.003103	0.002323	1.334	0.182296	0.99855	1.0077
Num. common groups	0.007395	1.007423	0.005746	1.287	0.198085	0.99614	1.0188
Num. photos together	-0.00229	0.997708	0.003061	-0.75	0.453433	0.99174	1.0037
Num. common friends	0.006295	1.006315	0.001621	3.883	0.000103	1.00312	1.0095

Notes: This table reports parameter estimates, hazard ratios, z-scores, confidence intervals and P-values for the Influence Cox proportional hazards model that estimate the impact of a user's tie attributes (tie strength and embeddedness) on his hazard to influence peers to adopt and on the hazard that his peers will spontaneously adopt, while controlling for individual influence and susceptibility associated with the individual attributes: age, gender or relationship status. Goodness of fit and model likelihood test statistics include: Concordance=0.832

(SE=0.009); Likelihood ratio test = 1797 on 67 degrees of freedom (p=0); Wald test statistic = 3489 on 67 degrees of freedom (p=0); Logrank statistic = 20640 on 67 degrees of freedom (p=0) ; Robust Logrank statistic = 411 (p=0)