

1-1-2010

Creating Social Contagion through Viral Product Design: A Randomized Trial of Peer Influence in Networks

Sinan Aral

NYU Stern School of Business, sinan@stern.nyu.edu

Dylan Walker

NYU Stern School of Business, dwalker@stern.nyu.edu

Follow this and additional works at: http://aisel.aisnet.org/icis2010_submissions

Recommended Citation

Aral, Sinan and Walker, Dylan, "Creating Social Contagion through Viral Product Design: A Randomized Trial of Peer Influence in Networks" (2010). *ICIS 2010 Proceedings*. Paper 44.
http://aisel.aisnet.org/icis2010_submissions/44

This material is brought to you by the International Conference on Information Systems (ICIS) at AIS Electronic Library (AISEL). It has been accepted for inclusion in ICIS 2010 Proceedings by an authorized administrator of AIS Electronic Library (AISEL). For more information, please contact elibrary@aisnet.org.

ICIS 2010 St. Louis

Creating Social Contagion through Viral Product Design: A Randomized Trial of Peer Influence in Networks

Completed Research Paper

Sinan Aral

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

Dylan Walker

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

Abstract

We examine how firms can create word of mouth peer influence and social contagion by incorporating viral features into their products. Word of mouth is generally considered to more effectively promote peer influence and contagion when it is personalized and active. Unfortunately, econometric identification of peer influence is non-trivial. We therefore use a randomized field experiment to test the effectiveness of passive-broadcast and active-personalized viral messaging capabilities in creating peer influence and social contagion among the 1.4 million friends of 9,687 experimental users. Surprisingly, we find that passive-broadcast viral messaging generates a 246% increase in local peer influence and social contagion, while adding active-personalized viral messaging only generates an additional 98% increase in contagion. Although active-personalized messaging is more effective per message and is correlated with more user engagement and product use, it is used less often and therefore generates less total peer adoption in the network than passive-broadcast messaging.

Keywords: Peer Influence, Social Contagion, Social Networks, Viral Marketing, Randomized Experiment.

Introduction

It is widely believed that social contagion and word of mouth (WOM) “buzz” about products drive product adoption and sales (Garber et al. 2004; Van den Bulte and Joshi 2007; Manchanda et al. 2008; Nam et al. 2009). Academic interest in the subject has recently exploded and managers are increasingly relying on “network” and “viral” marketing strategies to maximize returns to marketing investments (Hill et al. 2006). If firms can proactively manage WOM communication and viral buzz, they may be able to engineer the viral spread of their products to achieve wide spread adoption (Mayzlin 2006; Godes and Mayzlin 2009). Yet, although both managerial interest and academic research in this area are expanding dramatically, two dimensions critical to the success or failure of viral marketing efforts have been systematically understudied in the WOM literature – *viral product design* and the *econometric identification of peer influence*. We simultaneously address both of these topics by conducting a large scale randomized field experiment to test whether viral product features create peer influence and social contagion around a new commercial Facebook application. Our work applies lessons from the information systems literature and IT product design and testing to the problem of how to create effective proactive viral marketing online.

Viral product design – the process of explicitly engineering products so they are more likely to be shared amongst peers – has existed at least since the first chain letter was sent in 1888.¹ Today, many IT-enabled products and services firms attempt to design their products with features that make them likely to be virally shared among friends, family, colleagues, and acquaintances. For example, when Hotmail launched its free web-based email service in 1996, they designed a viral feature into the product by placing the following link and text as an embedded footer at the bottom of every email: “Get your private, free email at <http://www.hotmail.com>.” Each time a user sent an email, they passively and automatically advertised the service to the recipients. The information in the text of the footer facilitated the viral spread of awareness of the product, while the hypertext link facilitated the viral sharing of the product by providing each recipient a path to product adoption. Today the market is replete with products that have been engineered to ‘go viral’ with mixed results. Companies like Facebook technically enable users to ‘invite their friends’ to join the service through personalized referrals. When Google launched its Gmail service, personal referrals from other users were the only way one could obtain a Gmail account, creating an impression of exclusivity that Google hoped, when combined with a pervasive awareness campaign, would entice new users to demand the service from their friends. Today, when someone sends a Gmail message, an automated, pop-up hyperlink enables them to invite the recipient to join Gmail, but only if their email address is not already a Gmail address or if their non-Gmail address is not already known to be affiliated with an existing Gmail address, incorporating the sender’s judgment into targeting the referral process. Automated notifications are another viral design feature which inform peers of a user’s activity or use of a product. For example, automated notifications on LinkedIn build product awareness among peers and encourage users to return to the site to see what their contacts have been doing recently, while automated notifications sent from Facebook applications encourage friends of current application users to become aware of, interested in and to eventually adopt the application themselves.

Although some recent work examines firms’ proactive management of customer-to-customer communication for the purpose of creating WOM buzz and the viral spread of products, most of this work is focused on managing conversations about existing products rather than on proactively designing products to be viral (Mayzlin 2006; Godes and Mayzlin 2009). While a nascent literature addresses dimensions of viral product design that deal with inherent product characteristics (e.g. what makes a viral video ‘go viral’ or what makes a new story likely to be the ‘most emailed’ amongst peers) (Berger and Milkman 2009; Stephen and Berger 2009; Berger and Heath 2005; Phelps et al. 2004; Heath et al. 2001), less attention has been paid to firms’ use of viral product features to engineer virality (e.g. building messaging, hyperlinked embedding or automated referral and notification capabilities directly into products). Viral product features are fundamentally different in that they incorporate new modalities of product use into the design of a product to directly facilitate a) the sharing of the product or b) the peer-to-peer transfer of awareness about the product. To address this gap between theory and practice, we empirically evaluate the peer influence and social contagion effects of incorporating viral features into a product’s design.

Evaluating the effects of such product design decisions on social contagion is difficult however because the econometric identification of peer influence is non-trivial. As has been noted in economics, marketing and information systems literature, peer effects and WOM are typically endogenous (Manski 1993; Godes and Mayzlin

¹ This earliest known example of a chain letter seems to have originated as a part of a philanthropy effort initiated by four women in New Hampshire: <http://www.silcom.com/~barnowl/chain-letter/evolution.html>.

2004, 2009; Aral et al. 2009). While the spread of product adoption may correlate with WOM interactions between peers, simultaneity in the relationship between sales and WOM (Godes and Mayzlin 2004), bias from omitted variables such as advertising or inherent product quality (Van den Bulte and Lilien 2001; Godes and Mayzlin 2004), homophily in the preferences of linked peers (Aral et al. 2009), population heterogeneity (Thirtle and Ruttan 1987; Bemmaor 1994; Van den Bulte and Lilien 2001), truncation of observed data (Van den Bulte and Iyengar 2010), and other exogenous contextual effects (Manski 1993) could also explain such correlations. Although several identification strategies have been proposed, it is widely believed that randomized trials are the most effective way to obtain unbiased estimates of peer effects.

We therefore designed and conducted a randomized field experiment testing the effectiveness of two of the most widely used viral product features – personalized referral and automated notification – in creating peer influence and social contagion among the 1.4 million friends of 9,687 experimental users of Facebook.com. This is, to our knowledge, the largest randomized trial of peer influence ever conducted. The experiment utilizes a customized commercial Facebook application to observe user behavior, communications traffic and the peer influence effects of randomly enabled viral messaging capabilities on application diffusion in the local networks of experimental and control population users. By enabling and disabling the different viral features among randomly selected application users, we are able to obtain relatively unbiased causal estimates of the impact of ‘turning on’ a given viral feature on the adoption rates of peers in the local networks of adopters. As we record detailed click stream data we are also able to explore the underlying mechanisms by which viral features inspire adoption among peers.

Results show that viral product design features generate econometrically identifiable peer influence and social contagion effects. Features that require more activity on the part of the user and are more personalized to recipients create greater marginal increases in the likelihood of adoption per message, but generate fewer total messages, creating countervailing effects on peer influence. On average, passive-broadcast viral messaging capabilities, which are less personalized but also require less user effort, generate a 246% increase in local peer influence and contagion effects over a baseline model in which viral messaging is disabled. Adding active-personalized viral messaging capabilities, which are more personalized but require more effort, generates an additional 98% increase in peer influence and contagion above the passive-broadcast model. Analysis shows network externalities drive a viral feedback loop that accelerates contagion. These results shed light on how viral products can be designed to generate social contagion and how randomized trials can be used to identify peer influence effects in social networks.

Theory and Literature

WOM, Peer Influence and Viral Marketing

Early work by Katz and Lazarsfeld (1955) inspired great interest in how WOM can drive consumer choice and public opinion, and studies by Coleman et al. (1966), Griliches (1957), Arndt (1967) and Engel et al. (1969) corroborated the importance of peer influence in the diffusion of medical innovations, hybrid corn, food products and diagnostic devices for automobiles. Geographic correlations in aggregate sales data over time support inferences about the importance of WOM and the coefficient of imitation for product diffusion (Garber et al. 2004; Bell and Song 2007; Van den Bulte and Stremersch 2004; Manchanda et al. 2008). The importance of awareness and influence propagation through WOM communication are also supported by survey data on individuals’ participation in WOM behaviors (Bowman and Narayandas 2001). These data demonstrate correlations among personal recommendations (Reingen and Kernan 1986; Brown and Reingen 1987; Bowman and Narayandas 2001), online buzz (Godes and Mayzlin 2004), social network referrals (Reingen and Kernan 1986) and the diffusion of products and services from piano tuning to medical devices (Brown and Reingen 1987; Van den Bulte and Lilien 2001).

Evidence on the importance of WOM and its correlation with product sales and diffusion have led researchers to examine how firms might create broad, systematic propagation of WOM through consumer populations. Godes and Mayzlin (2009) refer to “endogenous WOM” and “exogenous WOM” to distinguish naturally occurring conversations among consumers from WOM communications ‘created as a result of firms’ actions.’ Evidence from their field test demonstrates the effectiveness of firm initiated buzz marketing, in which paid “agents” spread the word about products, generating exogenous WOM conversations where “none would have naturally occurred otherwise.” (Godes and Mayzlin 2009: 721) Researchers also examine advertising strategies that target those individuals most likely to propagate organic WOM most broadly. A long line of research suggests “influentials” drive product diffusion (Katz and Lazarsfeld 1955; Merton 1968; Gladwell 2000), although more recent simulation

studies suggest that cascades of influence are instead driven by “a critical mass of easily influenced individuals” (Watts and Dodds 2007). Influentials are identified by their persuasiveness, expertise, and the size and structure of their social networks (Reingen et al. 1984; Gladwell 2000; Goldenberg et al. 2001, 2009; Aral et al. 2009). Once a firm identifies whom to target, incentives to spread the word become critical and several studies address optimization of profitable referrals (Biyalogorsky et al. 2001; Libai et al. 2003; Ryu and Feick 2007).

Conspicuously absent from this large literature on viral marketing is work on viral product design. As Berger and Milkman (2009: 5) note “macro explanations for diffusion ... tend to ignore how individual level processes influence what gets shared ... Focusing on network structure ... and on the influence of special people provides little insight into why certain cultural items become viral while others do not ... Brown and Reingen (1987) note that “an enhanced understanding of social influence processes in consumer behavior may simply be obtained by examining which products or services consumers are more likely to “talk about.” (p.361), yet little empirical work has answered this call.” A small but growing literature has begun to examine the characteristics of content that make certain products viral. For example, Berger and Milkman (2009) find that awe-inspiring news stories that are practically useful, surprising, positive or affect-laden are more likely to make it into the New York Times “most emailed articles” list, and Heath et al.(2001) show that disgusting urban legends are more likely to be shared. This work extends a much larger and more general literature on the characteristics of products or innovations that influence collective adoption or diffusion (e.g. Rogers 2003). We complement and extend this work by proposing that product design features that enable and facilitate sharing and peer influence also contribute to product virality.

Viral Product Design

Can firms engineer products so they are more likely to be shared among peers? If so, which product features are most effective in inducing peer-to-peer influence in product adoption? We conceptualize viral product design – the process of explicitly engineering products so they are more likely to be shared amongst peers – as encompassing the incorporation of specific product characteristics and features into a product’s design to generate peer-to-peer influence in its adoption process. A product’s viral characteristics are fundamentally about its content and the psychological effects content can have on a user’s desire to share it with others (Berger and Milkman 2009; Stephen and Berger 2009; Berger and Heath 2005; Phelps et al. 2004; Heath et al. 2001). A product’s viral features on the other hand concern modalities of use with respect to sharing. Products may enable communication between users, generate automated notifications of each other’s activities, facilitate automated personalized invitations to peers or enable hypertext embedding of the product on publicly available websites and weblogs.

Enabled by both front end user interface features and backend database management technologies, *personalized referral* features enable users to select their friends or contacts from a list and to then invite them to join the service with the option to attach a personalized message to the invitation. Facebook users can invite their friends to join the social networking site and users can also invite their friends to download and use applications developed for use on the site. Gmail users can invite others to use Gmail and several other examples of this type of personalized invitation service exist in the marketplace. These services can be automated or require user initiation, can include a generic message or can be personalized, and can target contacts who are not currently users of the product or can alternatively target current users of a platform such as Facebook to adopt new platform specific applications.

Another feature is the *automated notification* triggered by user activity. When a user engages the product or takes an action which triggers the product to change the user’s status, these changes can be broadcast as notifications to the user’s contacts (whether or not their contacts are current users). For example, when a user of LinkedIn.com joins a new group, changes their profile information, connects to a new contact or takes a new job, their contacts are informed via email about the activity. The mobile geographic location service company FourSquare.com notifies users when their friends are nearby based on location tracking services that use mobile phone data. Facebook notifies friends when a user adopts a new application or achieves some application mile-stone and generally makes users aware of their friends’ various activities. Notifications such as these build awareness among friends of new activities or products a user is adopting or engaging with and can persuade peers to adopt these activities or products if the messages are persuasive of if knowledge of their friend’s engagement with the product is itself persuasive.

We classify viral features along two dimensions that affect awareness, sharing and preferences: *activity* and *personalization*. Activity describes the degree to which users must actively initiate the viral feature and ranges from ‘active’ to ‘passive.’ Active viral features require a user to actively choose to engage in sharing or interacting with other users or peers. For example, choosing a subset of one’s friends to invite to use a product is an active choice on

the part of the user. Typically, when invites are sent by users to their peers suggesting they adopt a product, the user actively chooses which friends to invite and what type of message to send to them as part of the invitation – each of these actions requires the user’s judgment and active participation. Passive viral features on the other hand are those that generate automated actions on behalf of the user without requiring the active choice of the user to engage those actions. For example, features that automatically monitor and broadcast a user’s geographical locations to peers without any actions taken by the users themselves (e.g. FourSquare.com), or those that automatically notify peers of a user’s activity with regard to the product without any active choice on the part of the user (e.g. Facebook notifications), are passive in that users’ judgment and active participation are not required to initiate the notification. The space of viral features ranges continuously on the activity dimension from completely passive to highly active. Personalization on the other hand describes the degree to which the output of the viral feature is personalized to each specific peer or more generally aimed at anyone, ranging from ‘broadcast’ to ‘personalized.’ Broadcast features enable engagement with the general population of possible consumers and are not directed specifically toward users’ personal contacts, while personalized features enable tailored engagement toward specific peers. For example, the hypertext embedding features in YouTube videos, Slide.com slideshows and RockYou widgets target any potential customer who sees them on the Web whether they are a friend or acquaintance of the user who posted them or not (features we label ‘generalized hypertext embedding’), while Facebook notifications are targeted only at a user’s personal social network. Personal referrals are more personalized than Facebook notifications because the user chooses a subset of their social network to whom the referral is sent. Referrals can be even more personalized if the user chooses to attach a personal note to the referral. The space of viral features ranges continuously along the personalization dimension from completely untargeted and not customized (broadcast) to individually targeted and tailored (personalized) (see Figure 1).

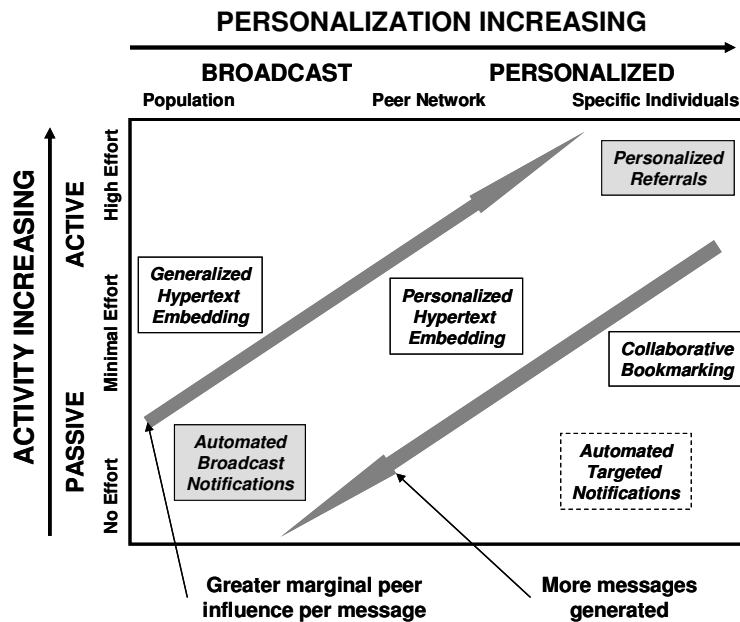


Figure 1: The Viral Product Feature Space

The viral product feature space has theoretically grounded implications for the likely effects of a given viral feature. These effects are denoted by the arrows that describe the gradients in the space pointing upward and to the right and downward and to the left. We argue that as features move upward and to the right in the viral product space toward active-personalized features like personal referrals and away from passive-broadcast features like automated broadcast notifications, their marginal effectiveness in creating peer influence and social contagion (inspiring others to adopt the product) will increase. Proactive invitations take time and energy to initiate and users’ must be aware they exist to use them, while automated notifications are simply generated by our online activities in the course of our normal behavior. However, when individuals take the time and effort to proactively choose to share information about products and services with their friends, they tend to choose to activate their strong tie relationships (Frenzen and Nakamoto 1993; Stephen and Lehmann 2009). Strong ties exhibit greater homophily (McPherson et al. 2001; Jackson 2008), greater pressure for conformity (Coleman 1988), and deeper knowledge about one another. We

simply know more about the preferences of our close friends and colleagues than we do about our acquaintances. We tend to trust information from close “trusted” sources more (typically our strong embedded ties) (Uzzi 1996), and we tend to respond more often to them out of a feeling of responsibility and reciprocity (Emerson 1962). In addition, the personalization of messages tends to make them much more effective, especially in online environments in which we are constantly bombarded with irrelevant information (Tam and Ho 2005; Dijkstra 2008). For these reasons it is likely that more personalized viral messages are more likely to be persuasive and effective even though they are less likely to be prevalent.

In summary, viral features range along the activity dimension of the viral product feature space from active to passive and along the personalization dimension from personalized to broadcast. Active-personalized viral features are generally considered to be more effective at promoting product contagion because personalized targeted messages are thought to be more persuasive. On the other hand, they require more effort which may curtail their use. The relative effectiveness of these viral features is ultimately an empirical question. We therefore designed and conducted a randomized experiment to estimate the effects of these viral product design features on product adoption and use.

Empirical Methods

Identification of Peer Influence in Social Networks

In order to accurately assess the impact of viral marketing or viral product design on product adoption and diffusion, it is important to seek econometrically identified parameter estimates of peer influence in the study population. Several sources of bias in both cross sectional and longitudinal data on interactions and outcomes among peers, including contextual and correlated effects, can confound assessments of peer influence and social contagion. First, WOM and product adoption may simply be simultaneously determined. If WOM is a function of sales and sales is simultaneously a function of WOM, reduced form estimates of one or the other relationship can be biased. As Godes and Mayzlin (2004:546) observe “high WOM today does not necessarily mean higher sales tomorrow. It may just be that the firm had high sales yesterday.” Even in longitudinal studies, identification of parameter estimates of these types of simultaneous relationships is difficult without exogenous instrumental variables or some other approach (Angrist et al.1996; Greene 2003). For this reason, Godes and Mayzlin (2004) are “careful to ... avoid any suggestions of causality” in their interpretations of parameter estimates. Second, omitted variables such as inherent product quality or common contextual effects, such as mass media exposure or marketing, can lead to incorrect estimates of peer influence. Van den Bulte and Lilien (2001) show that in Coleman et al’s (1966) seminal study of the diffusion of medical innovations, omission of data on marketing efforts led to overestimates of peer influence and social contagion. In fact, they find that when marketing is controlled for “contagion effects disappear.” Third, homophily can explain the clustering of product adoption decisions among socially connected peers (Aral et al. 2009). Individuals tend to choose friends with similar preferences (McPherson et al. 2001; Jackson 2008), and this assortativity on preferences can create clustering in product adoption decisions among peers even if these decisions are solely the result of individuals’ preferences for the product rather than peer influence. Finally, if decision relevant factors change over time in heterogeneous populations of consumers, the positive relationship between the likelihood of adoption and the prevalence of prior adoption in one’s local network – typically interpreted as evidence of social contagion – can be spurious (Van den Bulte and Lilien 2001). For example if a product’s price drops over time and customers’ willingness to pay is normally distributed, the acceleration of the adoption rate caused by a drop in price will also correlate with the increasing prevalence of adopters in consumers local networks over time.

Several approaches to the econometric identification of peer effects have been proposed in economics, sociology, marketing and information systems literatures including peer effects models, actor oriented models, instrumental variables methods based on natural experiments, dynamic matched sample estimation, structural models and ad hoc approaches based on specific data characteristics. Peer effects models and extended spatial autoregressive models capitalize on the idea that when local groups vary in size or structure, deviations from group means or particular structural configurations can, under certain assumptions, be identified using instrumental variables based on mean deviations or structural differences (Frank and Strauss 1986; Bramoulle et al. 2009; Kelejian and Prucha 1998; Lee 2003, 2007; Oestreicher-Singer and Sundararajan 2008). Actor oriented models characterize the co-evolution of network structure and behavior by modeling micro decisions that simultaneously maximize behavioral and network utility functions, and estimate continuous time Markov models on panel network data using MCMC or other

simulated method of moments techniques (Snijders and Baerveldt 2003; Snijders et al. 2006). Instrumental variables based on natural experiments utilize ‘exogenous’ changes in some individuals’ utility or behavior to identify their influence on peers (Tucker 2008; Brock and Durlauf 2001; Sacerdote 2001). Dynamic matched sample estimation evaluates differences between matched samples conditioned on a vector of observable characteristics, behaviors and attributes to recover influence estimates that account for the homophily that may make behaviors cluster in networks even if no peer influence exists (Rosenbaum and Rubin 1983; Hill et al. 2006; Aral et al. 2009). Structural models make assumptions about the specific form of the utility function that governs consumer choice, deriving identification conditions from those assumptions. Finally, ad hoc methods use the directionality of ties and changes in behaviors over time to corroborate causal interpretations of data (Christakis and Fowler 2007).

Yet, although each of these approaches provides improvements over traditional statistical methods, they have important weaknesses. Assumptions justifying the exogeneity of instrumental variables based on group mean deviations or network structure are debated. Actor oriented models do not converge easily on networks of greater than a few hundred nodes and characterize the proportional contributions of link formation and influence to observed outcomes without establishing causality per se. Instrumental variables based on natural experiments are rarely completely exogenous typically because relationships reflect social choices that reveal individuals’ preferences. Dynamic matched sample estimation, although applicable to very large datasets, requires a substantial amount of data in the vector of observable characteristics to create robust matches. As the correlation of peer preferences can come from unobservables, this method can only really bound influence estimates from above (Aral et al. 2009). Finally, it is difficult to establish the robustness of ad hoc methods with authority.

Randomized trials are widely believed to be the most effective way to obtain unbiased estimates of peer effects and the logic of randomization is quite simple (Falk and Heckman 2009). If we are interested in estimating the expected average effect of a treatment on a population of individuals, we cannot observe the expected outcome for a subject in the treatment group had she not been treated. Since in reality most individuals exposed to a treatment typically differ from those who are not, comparing the treated to the untreated without random assignment of the treatment creates a selection bias that reflects differences in the potential untreated outcomes of treatment and comparison groups. Randomization solves this problem because individuals assigned to the treatment and control groups differ in expectation only through their exposure to the treatment (Duflo et al. 2006). Thus, if the potential outcomes of an individual are also unrelated to the treatment status of any other individual, we can estimate the causal parameter of interest for the treatment (this is known as the “Stable Unit Treatment Assumption,” see Angrist et al. 1996). If the randomized trial is correctly designed and implemented, it can be shown that simple OLS estimation provides unbiased estimates of the treatment that are internally valid (Duflo et al. 2006: 8). We designed our experiment to ensure randomized treatment assignment and took great care to establish that the potential outcomes of any individual were unrelated to the treatment status of any other individual.

Experimental Design and Procedures

We partnered with a firm that develops commercial applications hosted on Facebook.com to elicit data on the peer influence effects of enabling viral features. The application we studied allows users to share information and opinions about movies, actors, directors and the film industry in general. The firm designed multiple experimental versions of the application in which personalized invitations and notifications were either enabled or disabled, and randomly assigned adopting users to various experimental and control conditions. When a user adopted the application, they were randomly assigned to one of the two treatment conditions or the baseline control condition, and the application collected their personal attributes and preferences from their Facebook profiles, as well as data on their social networks and the personal attributes and preferences of their network neighbors.²

The basic experimental design enabled experimental group users to use passive-broadcast and active-personalized viral messaging capabilities to exchange viral messages with their neighbors, while disabling these features for the baseline control group. The application then recorded data on the use of these viral features by experimental group users, as well as click stream data on recipient responses to viral messages, and their subsequent adoption and use of the application for all neighbors of experimental and control group users. When an individual adopted as a result of peer influence, their treatment status was also randomized to ensure that the Stable Unit Treatment Value Assumption held (Angrist et al. 1996). This facilitated analysis of the average treatment effect of enabling viral messaging capabilities on peer adoption and network propagation and allowed detailed analysis of the relative

² While Facebook allows blocking access to profile data less than 2% of users alter default privacy settings (Gross et al. 2005).

effectiveness of different viral messaging channels, as well as exploration of the mechanisms by which a particular viral channel influenced recipient behavior. The two primary viral features that were implemented in the application and enabled or disabled for different experimental treatment conditions are described below.

When enabled, *notifications* are generated automatically when an application user performs certain actions within the application, such as declaring a favorite movie or writing a movie review. When notifications are generated, they are distributed to a random subset of an application user's peers and displayed in a status bar at the bottom of the peers' Facebook environment. When a peer clicks on the notification, they are taken to an application canvas page where they are given the option to install the application. As no explicit action is necessary above and beyond the typical use of the application, notifications are classified as low effort on the activity dimension of the viral feature space. Furthermore, because notifications are randomly distributed to a Facebook user's peers and are not accompanied by a personal message, they are classified as low personalization (broadcast) in the viral feature space.

When enabled, *invites* allow an application user to send a personalized invitation to their Facebook peers, inviting them to install the application. A peer then receives the invitation in their Facebook inbox and may click on a referral link contained within the invitation. If they do, they are taken to the application canvas page where they are given the opportunity to install the application. As each invite requires a conscious and deliberate action on the part of the application user above and beyond typical application use, they are classified as higher effort (activity) than notifications in the viral feature space. Furthermore, because invites are targeted to specific Facebook peers and allow the inclusion of a personalized message, they are classified as higher in personalization than notifications.

The experimental design consisted of three treatment groups into which adopters of the application were randomly assigned: baseline, passive-broadcast, active-personalized. Users assigned to the baseline group received a version of the application in which both notifications and invites were disabled. In the passive-broadcast group (passive), only notifications were enabled. In the active-personalized group (active), both notifications and invites were enabled. There were no other differences between baseline, passive and active applications. Throughout the experiment, each adopter was randomly assigned to a treatment group according to the following proportions: 5% to the Baseline Group, 47.5% to the Passive-Broadcast Group and 47.5% to the Active-Personalized Group. Detailed logs of user activity, adoption times, viral feature use, peer response, and user and peer profile data were recorded.

At the launch of the experiment, an advertising campaign was designed in collaboration with a second Facebook advertising firm to recruit a population of application users. The advertising campaign was designed to reach a representative audience of Facebook users and advertisements were displayed to users through advertising space within other Facebook applications. The campaign was conducted in three waves throughout the duration of the experiment and cost a total of \$6000 to recruit 9687 usable experimental subjects, or 62 cents per recruit³. The number of impressions, clicks, and install responses to the recruitment campaign are displayed in Table 1.

Wave	Impressions	Clicks	Advertising Related Installs	Installs
1 (Day 0)	18,264,600	12,334	3,072	3,714
2 (Day 15)	20,912,880	25,709	2,619	3,474
3 (Day 20)	19,957,640	7,624	3,219	4,039
Total	59,135,120	45,667	8,910	11,227

Measurement, Analysis and Results

Data and Descriptive Statistics

The randomized experiment was conducted over a 44 day period during which 9687 users adopted the application with 405 users randomly assigned to the baseline control group, 4600 users randomly assigned to the passive-broadcast treatment group, and 4682 users randomly assigned to the active-personalized treatment group. Users in these groups collectively had 1.4M distinct peers in their local social networks and sent a total of 70,140 viral messages to their peers, resulting in 992 peer adoptions and 682 peer adoptions in direct response to viral messages.

³ The cost per recruited user is several times smaller than the cost-per-user associated with recruitment for lab-based experiments. The low cost of recruitment makes online experiments an excellent source of experimental data.

Three main observations arise from consideration of the summary statistics of the resultant data displayed in Table 2.

Table 2. Summary Statistics and Mean Comparisons of Active, Passive and Baseline Users

	1	2	3	4	5	6
	Baseline (N = 405)	Passive (N = 4600)	Active (N = 4682)	t-statistic (B-P)	t-statistic (B-A)	t-statistic (P-A)
	Mean (SD)	Mean (SD)	Mean (SD)	t-statistic (SE)	t-statistic (SE)	t-statistic (SE)
Age	31.51 (13.80)	30.81 (13.31)	29.94 (13.27)	.46 (13.35)	1.03 (13.31)	1.45 (13.24)
Gender (1=Male)	.25 (.44)	.33 (.47)	.32 (.47)	-1.57 (.47)	-1.42 (.46)	.40 (.47)
Degree [†]	171.79 (223.88)	170.25 (278.64)	166.97 (248.77)	.09 (275.13)	.32 (247.15)	.55 (263.82)
Number of Facebook Wall Posts	40.52 (79.89)	36.45 (94.16)	37.07 (246.76)	.46 (93.11)	.15 (238.20)	-.09 (188.31)
Number of Adopters in User's Local Network	.01 (.12)	0.07 (.35)	0.10 (.44)	-2.84*** (.34)	-3.60*** (.43)	-3.64*** (.40)
Percentage of Adopters in User's Local Network	.02 (.002)	.09 (.01)	.15 (.01)	-1.92* (.01)	-2.35** (.01)	-2.83*** (.01)
Maximum Diffusion Depth	.01 (.11)	.04 (.22)	.05 (.24)	-2.53* (.21)	-3.01*** (.24)	-1.98*** (.23)
Time to 1st Adopter	9.40 (9.71)	4.77 (8.04)	3.17 (6.72)	1.27 (8.07)	2.04** (6.77)	2.45*** (7.30)
Time to 2nd Adopter	---	5.23 (8.17)	4.43 (6.97)	---	---	0.58 (7.45)
Time to 3rd Adopter	---	5.29 (8.07)	3.04 (5.25)	---	---	1.08 (6.33)
Time to 4th Adopter	---	6 (5.83)	1.17 (1.12)	---	---	2.84*** (3.58)
Application Activity	3.17 (4.59)	4.17 (7.24)	4.56 (8.98)	-2.54** (7.08)	-2.89*** (8.73)	-2.20* (8.16)

***p<.001; **p<.05; *p<.10; † K-S Tests of Degree Distribution Differences: B-P: .04, p=.80, N.S.; B-A: .04, p=.79, N.S.; P-A: .01, p=.94, N.S.

First, assignment to control and treatment groups was clearly random with no significant mean or distributional differences between users assigned to baseline, passive-broadcast, and active-personalized treatments in terms of their age, gender, network degree (the number of Facebook friends), and their number of Facebook wall posts, confirming the integrity of the randomization procedure.

Second, while their demographics and Facebook activity patterns were the same, measures of peer response in the network neighborhoods of treated users differed significantly from control populations. T-tests show that the number and percentage of peer adopters in a user's local network are significantly higher for treated populations than for the baseline population. The number of peer adopters in a treated user's local network is roughly seven times greater for users that received the passive-broadcast treatment and ten times greater for users that received the active-personalized treatment as compared to that of users that received the baseline treatment. Similarly, the percentage of adopters in a user's local network is roughly 450% higher for users that received the passive-broadcast treatment and 750% higher for users that received the active-personalized treatment than in the networks of users that received the baseline treatment. The time to the first adopter is roughly 200% shorter for users that received the passive-broadcast treatment and roughly 300% shorter for users that received the active-personalized treatment as compared to users that received the baseline treatment. The average maximal distance in the social graph from a treated user to a peer adopter is approximately 360% greater for users that received the passive-broadcast treatment and approximately 450% greater for users that received the active-personalized treatment as compared to users that received the baseline treatment. These differences are all highly significant.

Finally, average application activity is roughly 130% higher for users that received the passive-broadcast treatment and 140% higher for users that received the active-personalized treatment than for users that received the baseline treatment. Two possible mechanisms could explain this increase in treated user activity. First, it could be that a more

viral application is simply more interesting and that this directly drives increased application use. Alternatively, it could be that application virality encourages peers of adopters to join them in application use, creating a positive feedback loop that inspires users to use the application more when their friends are using it. We examine these explanations in depth in the following sections.

Effects of Viral Product Design on Peer Influence and Social Contagion

Our main statistical approach uses hazard modeling, which is the standard technique for assessing contagion in economics, marketing, and sociology literatures (e.g. Van den Bulte and Lilien 2001; Iyengar et al. 2009, 2010; Nam et al. 2009). This approach represents the hazard of adoption of individual i at time t as a function of individual characteristics and social influence:

$$\lambda(t, x, w, y) = f(x_i(t)\gamma, \beta \sum_j w_{ij} y_j(t)),$$

where $\lambda(t)$ represents the baseline hazard of adoption; $x_i(t)$ is a vector of variables unrelated to social influence that affect i 's adoption decision; w_{it} is the social exposure of i to peer j ; $y_j(t)$ is the adoption status of peer j at time t ; and γ and β are parameters to be estimated.

Hazard rate models and binary choice models with duration dependence, which can be derived from utility theory and threshold based network models (Van den Bulte and Lilien 1999), are typically used to estimate such relationships (e.g. Van den Bulte and Lilien 2001; Manchanda et al. 2008). However, our circumstances require a slightly different approach as we are interested in estimating the treatment effects of randomly assigned viral features on the adoption of peers in the local networks of focal experimental and control users, rather than the effects of focal users' social environments on their own adoption decisions. We therefore estimate the peer effects of the treatment 'outward' from an individual to their peers rather than estimating the effects of an individual's social environment 'inward' on their own adoption hazard. Controlled "treatments" of each user's social environment are too complex and costly and observation of the diffusion of the product requires estimation of the hazards of the adoption of peers, and of the subsequent adoption of peers of peers. An 'inside-out' strategy estimating the effects of treatment on adoption in a user's social environment is therefore the most appropriate modeling approach.

Our approach compares the hazards of adoption in the social environments of users treated with passive and active viral applications to the hazards of adoption in the social environments of users treated with the baseline application. The analysis involves "multiple failure time" data that frequently arise in biomedical investigations in which multiple failures can occur for the same subject over time (Holmberg 2002). We want to estimate the hazard of multiple occurrences of peer adoption in the local networks of treated and untreated users as a function of their exposure to different viral features. In multiple failure time data, failure times are correlated within cluster (in our case within users' local networks), violating the independence of failure times assumption required in traditional survival analysis (Ezell et al. 2003). The simplest way to analyze multiple failure data is to examine 'time to first event' and several studies in the contagion literature take this approach (Iyengar et al. 2010). Other studies estimate the time to the first event and each subsequent event separately, which by construction assumes each sequential adoption event is equal and indistinguishable from the last (Anderson and Gill 1982). However, these specifications overlook potentially relevant data and fail to consider the cascading diffusion effects of multiple adoption events in a network, such as the presence of non-linear network effects or other non-linearities inherent in diffusion processes. We therefore employ a variance-corrected proportional hazards approach which adjusts the covariance matrix of the estimators in the model to account for the lack of independence among the multiple clustered failure times in the data, but allows the baseline hazards to vary by adoption event in order to account for the possibility that adoption hazards vary across stages of a diffusion process from first peer adopters to second peer adopters and so on.

Failure times in our adoption data are ordered, meaning there is a natural sequential ordering of event times such that the time of the first adoption in a local network by definition precedes the time of the second adoption and so on. If t_{ik} is the adoption time for the k th adoption in i 's network, adoption times are sequential such that $t_{ik} \geq t_{ik-1}$. As we observe time stamped adoption of the application in minutes and seconds, our data are ordered sequentially and no two events happen at the same time. As the social process of contagion can be affected by prior adoptions in a local network, for instance if network externalities are present, we assume that the baseline hazard function varies

over adoption occurrences, such that it differs from first adoption to second adoption to third adoption and so forth. We therefore estimate the following variance-corrected stratified proportional hazards model:

$$\lambda_k(t, X_{ki}) = \lambda_{0k}(t)e^{X_{ki}\beta}$$

where stratification occurs over the K adoption events, $\lambda_{0k}(t)$ represents the baseline hazard of the kth adoption event (*i*'s kth friend adopting); X_{ki} represents a vector of covariates affecting the adoption of *i*'s neighbors (including *i*'s viral treatment status (active, passive or baseline), a measure of *i*'s level of activity on the application (Application Activity), peer notifications sent (Notifications), and invites sent (Invites); and β is a vector of unknown parameters to be estimated (Prentice et al. 1981). We assume *i*'s kth friend does not adopt until their $k - 1$ friend adopts as this is the case for all our data. Therefore the conditional risk set at time t for event k consists of all subjects under observation at time t who have experienced a $k - 1$ adoption event (Cleves 1999). We estimate β using standard maximum likelihood estimation and adjust the covariance matrix to account for non-independence across individuals *i* using the following robust covariance matrix:

$$V = I^{-1}G'GI^{-1}$$

where *G* is a matrix of group efficient residuals. Results are presented in Table 3.

	1	2	3	4
	<i>Hazard Ratio</i>	<i>Hazard Ratio</i>	<i>Hazard Ratio</i>	<i>Hazard Ratio</i>
Viral State = Passive	3.46*** (1.18)	3.35*** (1.15)	2.50** (.86)	2.51** (.86)
Viral State = Active	4.44*** (1.64)	4.21*** (1.56)	3.33*** (1.24)	3.31*** (1.24)
Application Activity		1.02*** (.004)	1.02*** (.003)	1.02*** (.003)
Notifications			1.02*** (.002)	1.02*** (.002)
Invites				1.06** (.028)
Log Likelihood	-4694.359	-4631.795	-4544.845	-4542.577
χ^2 (d.f)	19.3*** (2)	57.4*** (3)	299*** (4)	308*** (5)
Observations	3929	3929	3929	3929

Notes: ***p<.001; **p<.05; *p<.10;

Table 3, Model 1 displays the average treatment effects of passive-broadcast and active-personalized viral treatments on peer influence and social contagion in the local networks of treated users above and beyond control group users who received the baseline application. Results indicate that users of the passive-broadcast application experienced a 246% increase in the rate of application adoption by peers compared to the baseline group, while active-personalized users experienced a 344% increase, demonstrating that inclusion of viral features creates peer influence and social contagion. Models 2-4 decompose the variance in local network adoption rates explained by these treatments by estimating how intermediate variables such as overall application activity, notifications and invites explain the resultant increases in peer adoption. For example Model 3 shows that a significant amount of the treatment effects are explained by correlated increases in users' use of the application and the viral messages their use generates. Users assigned to passive-broadcast and active-personalized applications use their applications more and send more messages (invites and notifications) that generate greater peer adoption in their local networks. Model 4 reveals that invites have a greater marginal impact on the adoption rate of peers than notifications. One additional personal invite increases the rate of peer adoption by 6%, while one additional notification increases the rate of peer adoption by 2% on average, confirming that more personalized active features have a greater marginal impact on the rate of peer adoption per message than passive broadcast features. Model 4 also demonstrates that variance explained by invitations only reduces the marginal effect of the active-personalized treatment variable and not the passive-broadcast treatment variable, which is what one would expect given the active-personalized treatment includes invitation capabilities while the passive-broadcast treatment does not.

The click stream data confirm these results. Table 4 displays the number of invitations and notifications sent, the responses to those messages that resulted in click through installations of the application and the resultant adoption rate per message (a proxy for the effectiveness of each type of message). Invitations, which require the most effort by the user and which are targeted specifically to recipients chosen by the user, are the least used but the most

effective per message in creating peer influence and social contagion. Notifications, which require the least effort and are automatically sent to randomly selected peers, generate the most messages, but are also the least effective per message. These results confirm the main finding findings of the study: viral product design features do in fact generate econometrically identifiable peer influence and social contagion effects. Features that require more activity on the part of the user and are more personalized to recipients create greater marginal increases in the likelihood of adoption per message, but also generate fewer messages resulting in less total peer adoption in the network.

	<i>Messages Sent</i>	<i>Adoptions via Click Through Installation</i>	<i>Adoption Rate (Marginal Impact)</i>
Invitations	160	16	.10
Notifications	69980	666	.001

Figures 2a) and 2b) plot the cumulative peer adoptions and the fractions of adopters in the local networks of baseline, passive and active treatment users, while 2d) plots the Kaplan-Meier survival estimates for baseline, passive and active treatments respectively. These graphs confirm that viral feature design has an economically significant impact on the diffusion of product adoption. Figure 2c) shows that treated users also use the application more than baseline users, which suggests that a positive feedback loop may be driving social contagion.

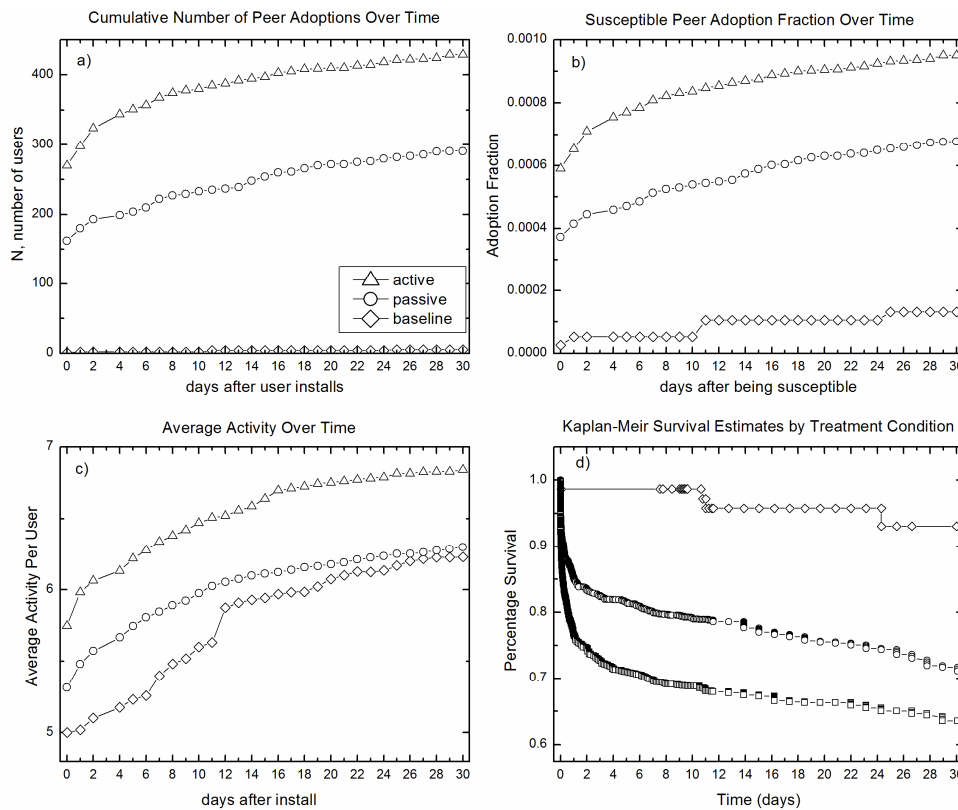


Figure 2: Plots a) the cumulative number of peer adoptions, b) the fraction of susceptible peer adopters, c) the average activity, and d) the Kaplan-Meier Survival Estimates over time for baseline, active and passive users.

Positive Feedback – Mechanisms Driving Social Contagion

Results of hazard models present strong evidence of social contagion effects caused by viral product design features. However, several social processes could underlie the dramatic impact of viral features. For instance, network externalities could generate a positive feedback loop of use and additional peer adoption (Van den Bulte and Stremersch 2004). It could also be that the features themselves may make the application more interesting and drive application use and peer adoption. We investigate alternate social explanations by examining relationships between

application features, activity and use, and application diffusion. Table 5 presents results estimating how these factors correlate with the number of peer adopters, diffusion depth and the user activity on the application.

Model 1 corroborates hazard rate estimates of social contagion. Controlling for user degree, passive-broadcast and active-personalized application users have significantly higher numbers of adopters in their local networks above and beyond the excluded baseline control group. Model 2 shows that these relationships hold when controlling for overall Facebook activity, which is expected since randomization ensures Facebook activity is constant across treatment and control groups. Model 3 demonstrates the importance of user application activity levels in explaining the number of peer adopters and shows that a primary channel through which treatments affect local peer adoption is through the viral messaging capabilities themselves. When notifications and invites variables are added to the regressions, they explain a significant amount of the variance originally attributed to the treatment effects, demonstrating that the viral features are actually driving treatment effects on peer adoption. Results in Model 3 also corroborate hazard model estimates, confirming that invitations have a higher marginal impact on peer adoption than notifications. Invitations are three times more effective per message than notifications in inspiring peer adoption on average. Models 4-6 confirm that the same pattern of results holds when estimating the depth of the contagion – how far out the product diffuses from control and treatment users. Active-personalized and passive-broadcast treatments significantly increase average diffusion depth, and these effects are again explained by application activity and the viral features themselves (Model 6). Finally, Models 7-10 examine how these factors explain application activity. Is it that the viral state of the application itself makes it more interesting, or rather is there a positive feedback loop that accelerates contagion? When the viral states are entered into the regression they significantly predict application activity in the expected directions and magnitudes across active-personalized and passive-broadcast application users when compared to the baseline (Model 7). However, when the number of peer adopters is entered into the analysis, these relationships disappear completely (Model 9). To confirm that it is not simply the utility from being “able to notify or invite friends” but rather friends’ actual adoption driving users’ activity, Model 10 controls for these factors and demonstrates that the more their peers adopt the application, the more users use the application. Evidence of a strong correlation between the number of adopter friends and application use suggest that a positive feedback loop exists that accelerates contagion.

Conclusion

We examined how firms can create word of mouth peer influence and social contagion by incorporating viral features into the design of their products. We presented a theory of viral product design based on a simple proposed space of viral product features that corresponds to predictions about how users will use different features and how effective they will be in generating peer influence and social contagion. We designed and conducted a randomized field experiment testing the effectiveness of passive-broadcast and active-personalized viral messaging capabilities in creating peer influence and social contagion among the 1.4 million friends of 9,687 experimental users of Facebook.com. Utilizing a customized commercial Facebook application to observe user behavior, communications traffic and the peer influence effects of different viral product design choices, we showed that viral product features can in fact generate econometrically identifiable peer influence and social contagion effects. Features that require more activity on the part of the user and are more personalized to recipients create greater marginal increases in the likelihood of adoption per message, but generate fewer total messages creating countervailing effects on peer influence. On average, passive-broadcast viral messaging capabilities, which are less personalized but also require less user effort, generated a 246% increase in local peer influence and contagion effects over a baseline model in which viral messaging is disabled. Adding active-personalized viral messaging capabilities, which are more personalized but require more user effort, generates an additional 98% increase in local peer influence and contagion effects over the passive-broadcast model. Analysis showed that initial peer adoptions in users’ local networks drive a viral feedback loop that accelerates contagion. The results shed light on how viral products can be designed to generate social contagion and how randomized trials can be used to identify peer influence in social networks.

Table 5. Drivers of Application Diffusion										
	1	2	3	4	5	6	7	8	9	10
	Number of Adopters	Number of Adopters	Number of Adopters	Diffusion Depth	Diffusion Depth	Diffusion Depth	Application Activity	Application Activity	Application Activity	Application Activity
	<i>Beta</i> (<i>SE</i>)	<i>Beta</i> (<i>SE</i>)	<i>Beta</i> (<i>SE</i>)	<i>Beta</i> (<i>SE</i>)	<i>Beta</i> (<i>SE</i>)	<i>Beta</i> (<i>SE</i>)	<i>Beta</i> (<i>SE</i>)	<i>Beta</i> (<i>SE</i>)	<i>Beta</i> (<i>SE</i>)	<i>Beta</i> (<i>SE</i>)
Viral State = Passive	.078** (.031)	.084** (.033)	.020 (.059)	.045** (.0178)	.048*** (.019)	.020 (.018)	.129* (.074)	.112 (.079)	.062 (.076)	-.037 (.074)
Viral State = Active	.119*** (.031)	.131*** (.032)	.059* (.030)	.057*** (.018)	.063*** (.019)	.033* (.018)	.190*** (.074)	.171** (.079)	.091 (.076)	-.006 (.074)
Degree	.0001*** (.00002)	.0001** (.00003)	.0001** (.00002)	.0001*** (.00001)	.00004** (.00002)	.00003** (.00001)	-.0001 (.0001)	-.0001 (.0001)	-.0002** (.0001)	-.0002** (.0001)
Facebook Activity		.019*** (.006)	.006 (.006)		.013*** (.004)	.007** (.004)		.054*** (.016)	.042*** (.015)	.026* (.014)
Application Activity			.061*** (.005)			.021*** (.003)				--
Notifications			.010*** (.0004)			.005*** (.0002)				.022*** (.001)
Invites			.035*** (.010)			-.003 (.006)				.055** (.024)
Number of Adopters									.607*** (.030)	.360*** (.031)
F Value (d.f.)	12.20*** (3)	11.18*** (4)	157.94*** (7)	9.36*** (3)	10.11*** (4)	85.13*** (7)	3.51*** (3)	4.87*** (4)	83.54*** (5)	128.92*** (7)
R ²	.006	.007	.16	.004	.007	.09	.002	.003	.07	.14
Observations	8910	5766	5766	6310	5766	5766	6310	5766	5766	5766

Notes: ***p<.001; **p<.05; *p<.10. These models are estimated with OLS regression.

References

- Andersen PK, RD Gill "Cox's Regression Model for Counting Processes: A Large Sample Study," *Annals of Statistics* 10, 1982, pp. 1100-20.
- Angrist JD, Imbens GW, & Rubin DB "Identification of Causal Effects Using Instrumental Variables", *Journal of the American Statistical Association* 91, 1996
- Aral, S., L. Muchnik, & A. Sundararajan "Distinguishing influence-based contagion from homophily-driven diffusion in dynamic networks," *P. Natl. Acad. Sci.* 106, 2009, pp. 21544-21549.
- Aral, S and Van Alstyne, M. "Networks, Information & Brokerage: The Diversity-Bandwidth Tradeoff", 2009. Available at SSRN: <http://ssrn.com/abstract=958158>.
- Arndt J "Role of Product-Related Conversations in the Diffusion of a New Product," *Journal of Marketing Research* 4(3), 1967, pp. 291-295.
- Bell D & Song S "Neighborhood effects and trial on the internet: Evidence from online grocery retailing," *Quantitative Marketing and Economics* 5(4), 2007, pp. 361-400.
- Bemmaor, AC "Modeling the diffusion of new durable goods: Word of mouth effect versus consumer heterogeneity," *Research Traditions in Marketing*, eds. G Laurent, GL Lilien, B Pras. Kluwer, Boston, 1994.
- Berger J & Milkman K "Social Transmission and Viral Culture," Working Paper, 2009, Wharton
- Berger J and Heath C "Idea habitats: How the prevalence of environmental cues influences the success of ideas," *Cognitive Science* 29 (2), 2005, pp. 195-221.
- Biyalogorsky E, Gerstner E, & Libai B (2008) "Customer Referral Management: Optimal Reward Programs,". *Marketing Science* 20(1), 2001, pp. 82-95
- Bowman D, D Narayandas "Managing customer-initiated contacts with manufacturers: The impact of share category requirements and word of mouth behavior," *Journal of Marketing Research* 38, 2001, pp. 291-297.
- Bramoulle Y, Djebbari H, & Fortin B "Identification of Peer Effects through Social Networks," *Journal of Econometrics* 150(1), 2009, pp. 41-55.
- Brock WA & Durlauf SN "Discrete Choice with Social Interactions," *The Review of Economic Studies* 68(2), 2001, pp. 235-260.
- Brown JJ & Reingen PH "Social Ties and Word-of-Mouth Referral Behavior," *The Journal of Consumer Research* 14(3), 1987, pp. 350-362.
- Christakis NA & Fowler JH "The Spread of Obesity in a Large Social Network over 32 Years,". *N Engl J Med* 357(4), 2007, pp. 370-379.
- Cleves M "Analysis of multiple failure-time survival data," Stata Technical Bulletin STB-49, 1999
- Coleman JS "Introduction to Mathematical Sociology," New York: Free Press., 1964
- Coleman JS "Free Riders and Zealots: The Role of Social Networks," *Sociological Theory* 6(1), 1988, pp. 52-57.
- Coleman JS, E Katz, H Menzel "Medical Innovation: A Diffusion Study," Bobbs-Merrill, Indianapolis, 1966
- Dijkstra A "The Psychology of Tailoring-Ingredients in Computer-Tailored Persuasion," *Social and Personality Psychology Compass* 2(2), 2008, pp. 765-784.
- Duflo E, Glennerster R, & Kremer M "Using Randomization in Development Economics Research: A Toolkit", *Center for Economic Policy Research Discussion Paper Series no. 6059*, 2006
- Emerson RM "Power-Dependence Relations," *American Sociological Review* 27(1), 1962, pp. 31-41.
- Engel JS, Blackwell RD, Kegerreis RJ "How information is used to adopt and innovation," *Journal of Advertising Research* 9(4), 1969, pp. 3-8.
- Ezell ME, Land KC, & Cohen LE "Modeling Multiple Failure Time Data: A Survey of Variance-Corrected Proportional Hazards Models with Empirical Applications to Arrest Data," *Sociological Methodology* 33(1), 2003, pp. 111 - 167.
- Falk A & Heckman JJ "Lab Experiments Are a Major Source of Knowledge in the Social Sciences," *Science* 326(5952), 2009, pp. 535 - 538.
- Frank O & Strauss D "Markov Graphs," *Journal of the American Statistical Association* 81(395), 1986, pp. 832-842.
- Frenzen J & Nakamoto K "Structure, cooperation, and the flow of market information," *Journal of Consumer Research* 20(3), 1993, pp. 360-375.
- Garber TJ, Goldenberg B, Libai E and Muller E "From density to destiny: Using spatial dimension of sales data for early prediction of new product success," *Marketing Science* 23(3), 2004, pp. 419-428.
- Gladwell, M "The Tipping Point: How little things can make a big difference," Little Brown, NY, 2000
- Godes D & Mayzlin D "Using Online Conversations to Study Word-of-Mouth Communication," *Marketing Science* 23(4), 2004, pp. 545-560.

- Godes D & Mayzlin D "Firm-Created Word-of-Mouth Communication: Evidence from a Field Test," *Marketing Science* 28(4), 2009, pp. 721-739.
- Goldenberg J, Han S, Lehmann DR, & Hong JW "The Role of Hubs in the Adoption Process," *Journal of Marketing* 73(2), 2009, pp. 1-13.
- Goldenberg J, Libai B, & Muller E "Talk of the Network: A Complex Systems Look at the Underlying Process of Word-of-Mouth," *Marketing Letters*, 2001, pp. 211-223.
- Greene WH "Econometric Analysis" Prentice Hall, 4 Ed., 2003
- Griliches Z "Hybrid Corn: An exploration in the economics of technological change," *Econometrica* 25:4, 1957, pp. 501-522
- Gross R., Acquisti A, & Heinz J "Information Revelation and Privacy in Online Social Networks", *Workshop on Privacy in Proceedings of the 2005 ACM Workshop on Privacy*, 2005
- Heath C, Bell C, and E Sternberg "Emotional selection in Memes: The case of urban legends," *Journal of Personality and Social Psychology* 81(6), 2001, pp. 1028-1041.
- Hill S, Provost F, & Volinsky C "Network-Based Marketing: Identifying Likely Adopters via Consumer Networks," *Statist. Sci.* 21(2), 2006, pp. 256-276.
- Holmberg L, Bill-Axelsson A, Helgesen F, Salo JO, Folmerz P, Haggman M, Andersson S, Spangberg A, Busch C, Nordling S, Palmgren J, Adami H, Johansson J, Norlen BJ, "A Randomized Trial Comparing Radical Prostatectomy with Watchful Waiting in Early Prostate Cancer," *The New England Journal of Medicine* 347(11), 2002, pp. 781-789
- Iyengar R, Bulte CVd, & Valente TW "Opinion Leadership and Social Contagion in New Product Diffusion", *Wharton Working Paper*, 2010
- Iyengar R, Han S, & Gupta S "Do Friends Influence Purchases in a Social Network?", *Harvard Working Paper 09-123*, 2009
- Jackson MO "Average Distance, Diameter, and Clustering in Social Networks with Homophily," *Internet and Network Economics*, Springer, Berlin / Heidelberg, 2008, , Vol 5385/2008, pp. 4-11.
- Katz, E., P.F. Lazarsfeld "Personal Influence," Free Press, New York, 1955
- Kelejian HH & Prucha IR "The method of elimination and substitution in the GMM estimation of mixed regressive, spatial autoregressive models," *The Journal of Real Estate Finance and Economics* 17(1), 1998, pp. 99-121.
- Lee L-f "Best Spatial Two-Stage Least Squares Estimators for a Spatial Autoregressive Model with Autoregressive Disturbances," *Econometric Reviews* 22(4), 2003
- Lee L-f "Identification and estimation of econometric models with group interactions, contextual factors and fixed effects." *Journal of Econometrics* 140(2), 2007, pp. 333-374.
- Libai B, Biyalogorsky E, & Gerstner E "Setting Referral Fees in Affiliate Marketing," *Journal of Service Research* 5(4), 2003, pp. 303-315.
- Manchanda P, Xie Y, & Youn N "The Role of Targeted Communication and Contagion in Product Adoption," *Marketing Science* 27(6), 2008, pp. 961-976.
- Manski CF "Identification of Endogenous Social Effects: The Reflection Problem," *The Review of Economic Studies* 60(3), 1993, pp. 531-542.
- Manski CF "Identification Problems in the Social Sciences," *Sociological Methodology* 23, 1993, pp. 1-56.
- Mayzlin, D. "Promotional Chat on the Internet," *Marketing Science* 25(2), 2006, pp. 155-976.
- Merton, RK "Patterns of Influence: Local and Cosmopolitan Influentials," *Social Theory and Social Structure*, ed. Robert K. Merton, New York: Free Press, 1968, pp. 441-74.
- McPherson M, Smith-Lovin L, & Cook JM "Birds of a Feather: Homophily in Social Networks," *Annual Review of Sociology* 27, 2001, pp. 415-444.
- Nam, S, Manchanda, P and PK Chintagunta "The effects of service quality and word of mouth on consumer acquisition, retention and usage," Forthcoming in *Marketing Science*, 2009
- Oestreicher-Singer G & Sundararajan A "The Visible Hand of Social Networks in Electronic Markets," *Working Paper*, 2008, Available at <http://ssrn.com/abstract=1268516>
- Phelps, JE, L Lewis, L, Mobilo, D Perry, N Raman "Viral Marketing or Electronic Word-of-Mouth Advertising: Examining Consumer Responses and Motivations to Pass Along Email," *Journal of Advertising Research* 44, 2004, pp. 333-348
- Prentice, RL, BJ Williams, AV Peterson "On the regression analysis of multivariate failure time data," *Biometrika* 79, 1981, pp. 495-512.
- Reingen P, B Foster, JJ Brown, S Seidman "Brand congruence in interpersonal relations: A social network analysis," *Journal of Consumer Research* 11, 1984, pp. 1-26.

- Reingen P, Kernan, JB “Analysis of referral networks in marketing: methods and illustration,” *Journal of Marketing Research* 23:4, 1986, pp. 370-378.
- Ryu G & Feick L “A Penny for Your Thoughts: Referral Reward Programs and Referral Likelihood,” *Journal of Marketing* 71(1), 2007, pp. 84-94.
- Sacerdote B “Peer Effects with Random Assignment: Results for Dartmouth Roommates,” *Quarterly Journal of Economics* 116(2), 2001, pp. 681-704.
- Snijders T, Steglich C, & Schweinberger M “Modeling the co-evolution of networks and behavior,” *Longitudinal models in the behavioral and related sciences*, eds Montfort Kv, Oud H, & Satorra A (Lawrence Erlbaum), 2006
- Snijders TAB & Baerveldt C “A Multilevel Network Study of the Effects of Delinquent Behavior on Friendship,” *Journal of Sociology* 27, 2003, pp. 123-151.
- Stephen AT & Berger JA “Creating Contagious: How Social Networks and Item Characteristics Combine to Spur Ongoing Consumption and Reinforce Social Epidemics,” *Wharton Working Paper*, 2009
- Stephen AT & Lehmann DR “Why Do People Transmit Word-of-Mouth? The Effects of Recipient and Relationship Characteristics on Transmission Behaviors,” *Wharton Working Paper*, 2009
- Tam KY & Ho SY “Web Personalization as a Persuasion Strategy: An Elaboration Likelihood Model Perspective,” *Information Systems Research* 16(3), 2005, pp. 271-291.
- Thirtle, CG, and Ruttan, VW “The role of demand and supply in the generation and diffusion of technical change,” Harwood Academic, 1987
- Tucker C “Identifying Formal and Informal Influence in Technology Adoption with Network Externalities,” *Management Science* 54(12), 2008, pp. 2024-2038.
- Uzzi B “The Sources and Consequences of Embeddedness for the Economic Performance of Organizations: The Network Effect,” *American Sociological Review* 61(4), 1996, pp. 674-698.
- Van den Bulte C & Iyengar R “Tricked by Truncation: Another Source of Spurious Social Contagion in New Product Adoption,” *Wharton Working Paper*, 2010
- Van den Bulte C & Joshi YV “New Product Diffusion with Influentials and Imitators,” *Marketing Science* 26(3), 2007, pp. 400-421.
- Van den Bulte C & Lilien GL “Medical Innovation Revisited: Social Contagion versus Marketing Effort,” *The American Journal of Sociology* 106(5), 2001, pp. 1409-1435.
- Van den Bulte C & Lilien GL “Integrating Models of Innovation Adoption: Social Contagion, Utility Maximization and Hazard Models,” *Working Paper* 99-021, 1999, Wharton School.
- Van den Bulte, C. and S. Stremersch “Social Contagion and Income Heterogeneity in New Product Diffusion: A Meta-Analytic Test,” *Marketing Science* 23(4), 2004, pp. 530-544.
- Watts DJ & Dodds PS “Influentials, Networks, and Public Opinion Formation,” *Journal of Consumer Research* 34, 2007