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

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

We examine how firms can create word of mouth peer influence and social contagion by incorporating viral features into the design of their products. Evaluating the effects of such product design decisions on social contagion is difficult because econometric identification of peer influence is non-trivial. Although several approaches have been proposed, it is widely believed that the most effective way to obtain unbiased estimates of peer effects is to conduct large-scale randomized trials of peer-to-peer communications intended to influence particular economic decisions, such as the decision to adopt a product. We therefore 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. 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. 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 user effort, generates an additional 98% increase in local peer influence and contagion effects over the passive-broadcast model. Analysis shows that initial peer adoptions in users’ local networks 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.

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

Draft: February, 2010

* We are grateful to Haim Mendelson, Lev Muchnik, Foster Provost and seminar participants at the Workshop on Information Systems Economics, MIT Sloan, NYU Stern and Wharton for invaluable discussions. Sinan Aral acknowledges generous financial support from the National Science Foundation (CAREER Award 0953832), IBM and the Institute for Innovation and Information Productivity.
1. 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. In this way, 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.

Marketing literature has a long tradition in studies of WOM communication, product adoption and diffusion and viral marketing, but is only recently focusing on how IT plays a role in these processes (Godes and Mayzlin 2004, Chevalier and Mayzlin 2006). On the other hand information systems literature and IT product and services firms are well aware of how the unique characteristics of IT enable both rapid prototyping and testing, and the precise study of human behavior. We combine lessons from both traditions to build and test a theory of how viral product design can create social contagion.

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 requesting donations for
education efforts in the Cumberlands region of New Hampshire in 1888.¹ Today, many IT-enabled prod-
ucts 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 follow-
ing 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 email’s 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. The intention was not only to spread awareness of Hotmail, but
to increase its rate of adoption through traffic over the hypertext link. Hotmail subsequently grew from
zero to 30 million users in 30 months and was acquired by Microsoft a year after its launch for $400 mil-
lion.

Hotmail is but one example. 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 re-
ferrals 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, Google uses a more sophisticated viral feature.
When someone sends a Gmail message, an automated, pop-up hyperlink enables them to invite the recipi-
ent to join Gmail, but only if their email address is not already a Gmail address or if their non-Gmail ad-
dress is not already known to be affiliated with an existing Gmail address. Unlike Hotmail, this feature
incorporates the sender’s judgment into targeting the referral process and uses backend database algo-

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. The details and text of this letter may be found at:
rithms to ensure referrals are not wasted on those who already use Gmail. On the other hand, it also requires users to expend more effort to activate the viral feature than in the case of Hotmail, creating indeterminacy around which feature is more likely to be effective.

*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. Increasingly, these viral features are also being incorporated into the design and delivery of IT-enabled services that complement non-IT related products such as automated referral and reminder features for flower delivery services like 1800Flowers or social news content sharing via the New York Times or Google’s Buzz service. We will discuss these and other viral design features (and why they might succeed or fail) in detail Section 2, but a cursory scan of the market for IT-enabled products and services reveals many viral design features incorporated into products.

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 are the dimensions of email stories that make them 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, Bell and Sternberg 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). Although this recent work examines how products’ inherent characteristics, such as their usefulness, topicality, prominence, positive valence and unexpectedness, can make them viral (Berger and Milkman 2009, Stephen and Berger 2009), viral product features are funda-
mentally 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 prod-
uct. To address this gap between theory and practice, we empirically evaluate the peer influence and so-
cial 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 2004, 2009, Aral et al 2009). While the spread of product adoption may corre-
late 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 contextu-
tial effects (Manski 1993) could also explain correlations between social network ties and the propensity
of peers of prior adopters to adopt a given product. These alternative explanations can account for a great
deal of what at first appears to be a contagious process (Aral et al 2009). Although several identification
strategies have been proposed (in literature we review below), it is widely believed that the most effective
way to obtain unbiased estimates of peer effects is to conduct large-scale randomized trials of peer-to-
peer communications intended to influence particular economic decisions, such as the decision to adopt a
product. As Godes and Mayzlin (2004: 558) note “it is very difficult to draw clean inferences of causality
with traditional econometrics … more work is needed to identify the causal link between WOM and fu-
ture sales.”

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. 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 on users’ online behaviors we are also able to explore in detail the underlying mechanisms by which a given viral feature inspires adoption among a users’ 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 completely disabled. Adding active-personalized viral messaging capabilities, which are more personalized but also require more effort, generates an additional 98% increase in local peer influence and contagion effects over and above the passive-broadcast model. Analysis shows that initial peer adoptions in users’ local networks 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.

2. Theory and Literature

2.1. 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 (1967), Grilliches (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 ag-
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 2004, Van den Bulte and Stremersch 2004, Manachnda et al 2008), and collocation studies suggest that WOM and interpersonal communication create brand preference congruity among women living in the same sorority (Reingen et al 1984) and correlations in the adoption of high yield variety (HYV) seeds among Indian farmers living in the same village (Foster and Rosenzweig 1995). Although inferences based on these types of aggregate observational data can be confounded by alternate explanations (Van den Bulte and Lilien 2001, Aral et al 2009), theoretical work supports the argument that people are indeed influenced by the opinions, actions and suggestions of their peers (Banerjee 1992, 1993). As information cascades from person to person through WOM communications, decisions based on information inferred from peer actions can create systematic “herding” behaviors that explain fads and bubbles, even around suboptimal outcomes (Bikhchandani et al 1991). The importance of awareness and influence propagation through WOM communication are also supported by more specific 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), referrals through social networks (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) Mayzlin (2006) provides theoretical support for the potential effectiveness of firm created WOM by showing that
equilibrium strategies exist in which firms would profit from posing as customers to generate favorable online WOM even when customers are aware they might be doing so.

Proactive firm efforts to create and optimize WOM are not only aimed at creating new online conversations or employing paid buzz agents to raise awareness. Researchers also examine advertising strategies that target those individuals most likely to propagate organic WOM most broadly. In this line of inquiry, the two main questions are whom to target and how to incentivize them to spread the message? 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 (Gladwell 2000, Goldenberg et al 2009). Literature on this latter type of “network” marketing considers how individuals’ positions in social network structure may enable them to actualize broad based diffusion through peer influence (e.g. Watts 2002, Iribarran and Moro 2009). This work privileges the importance of social hubs (Goldenberg et al 2009), examines how strong and weak ties and network size interact to affect message propagation (Goldenberg et al 2001) and studies how similarities within and across cohorts impact product diffusion (Reingen et al 1984, Aral et al 2009). Once a firm identifies whom to target, how to optimally incentivize them to spread the word becomes critical. In this domain, several studies address optimization of profitable referrals (Biyalogorsky et al 2001, Libai et al 2003, Ryu and Feick 2007). For example, Biyalogorsky et al (2001) show that firms should offer rewards for customer referrals only if people are not too demanding.

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. We suggest that viral product design, including the development of both product characteristics and product features, is an important area of future research on the virality of products with significant managerial relevance.

2.2. 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? Though a strong tradition links product design and marketing theory through techniques such as conjoint analysis and discrete choice modeling for optimal product feature (or attribute) selection (Kuhfeld et al 1994, Bloch 1995, Souder and Song 1997, Green et al 2001), this work is aimed squarely at analyzing individual-level preferences for products. Experimental studies in this vein test which “bundles of attributes” or product features best satisfy consumers’ preferences, create the greatest utility and ultimately maximize profit. Complementary investigations of product virality could improve product design and marketing by incorporating consideration of social contagion into our conceptualization of the consumption of a product by a socially connected population of consumers. If a product’s design can make it spread from user to user, this could have profound implications for the joint optimization of product design and marketing.

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 charac-
Creating Social Contagion through Viral Product Design

teristics 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. A nascent literature has begun to study which content characteristics are likely to inspire viral sharing and diffusion, and characteristics identified to date include usefulness, topicality, prominence, positive valence and unexpectedness (Berger and Milkman 2009, Stephen and Berger 2009, Berger and Heath 2005, Phelps et al 2004, Heath, Bell and Sternberg 2001). A product’s viral features on the other hand concern modalities of use with respect to sharing – how features enable and constrain a product’s use in relation to other consumers. 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. Several examples serve to characterize viral features in more depth.

**Personalized Referrals.** 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. For example, Plaxo, the contact management service company, provides IT-enabled features that automate the process of migrating, updating and maintaining contact information across multiple platforms. The system mines email address books and addresses found in emails to send contacts requests for updated information. As contacts of users may or may not be Plaxo users themselves, these requests serve to build awareness about the product among a users’ contacts and to invite them to join the service as well. Ning.com, a social network development platform provider encourages users to build their own social networks, enables them to invite their friends to the network and enables their friends to subsequently build their own networks around their niche interests. Niche networks in turn allow advertisers to target consumers with niche advertisements. 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 products or applications.

*Hypertext Embedding.* Although YouTube.com is primarily known for enabling the spread of viral videos that are passed from user to user due to the inherent viral nature of the video content (the video’s characteristics), the popular video streaming site also explicitly incorporates viral features into their product design. By allowing anyone to embed a video or video link into their own weblog, website or MySpace page, they enable and encourage uses of their product (and the video content their product delivers) that facilitate sharing and peer-to-peer transfers of awareness about the product. As friends, relatives, acquaintances and strangers watch videos embedded on websites and weblogs, they are consuming YouTube content. At the end of a video they are shown a piece of code that enables them to embed the video in their own website or weblog and a link that enables them to share the video personally via email with their friends. If they do not want to embed or share that particular video, other videos are recommended using collaborative filtering algorithms that analyze correlations in users’ preferences across video content. The *embedding feature* creates virality by enabling and encouraging users to share and promote YouTube content. The simplicity of the design feature makes sharing easy for anyone to accomplish and recommendation engines help suggest content that users are statistically more likely to like and share. Several other prominent products also utilize this design feature. For example, Slide.com and RockYou.com provide products and services that allow users to create and embed slideshows of pictures or other content on their websites, weblogs and social networking profile pages. As other users browse those items on the web, hyperlinks allow them to download the products or services themselves. In this way, embedding enables users to spread awareness and to provide a path to product adoption for other users.

*Automated Broadcast Notifications.* A third 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 in some way, these changes can be broadcast as notifications to the user’s contacts (whether
Creating Social Contagion through Viral Product Design

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 milestone 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 or if knowledge of their friend’s engagement with the product is itself persuasive.

These types of viral product features are likely to promote sharing, diffusion and contagion by affecting peer awareness and peer preference (Godes and Mayzlin 2009). Models of persuasion, such as the Elaboration Likelihood Model (Petty and Cacioppo 1986), typically conceptualize stages of information processing that lead to persuasion as consisting of attention, elaboration, and finally decision making or behavior (Bargh 2002), and the process of induction, whereby a user’s use of a product inspires a peer to adopt and use the product, involves making the peer aware of the product, enabling them to share the product and persuading them to change their expected utility such that they adopt.

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

---

2 By ‘share’ we mean provide a path to product adoption.
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 along 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.3 As is clear from this example, personalization encompasses both targeting and customization. Targeting specifies whether the feature is directed at the broad population of potential consumers, a subset of consumers like a current user’s social network, or a specific person (Hill et al 2006). Customization specifies whether the content of a feature’s engagement with the recipient can be tailored to a group of friends or a specific individual with a personalized message that is either actively written by the user or passively generated by the system (Tam and Ho 2005).

3 Generic messages can be replaced or complemented with personalized text. For example, an invitation sent from the Flixster Facebook application begins with the default text: “Hey - allow this app access so we can test our Movie Compatibility.” User’s sending the invitation may click the “add a personal message” button to customize the invitation.
The space of viral features ranges continuously along the personalization dimension from completely un-targeted and not customized (broadcast) to individually targeted and tailored (personalized).\(^4\)

When combined, the two dimensions describe a continuous space of viral product features that range from broadcast features generally aimed at anyone to personalized features targeted and tailored toward specific peers, and from active features that require active user engagement to passive features that generate automated actions on behalf of the user, and any convex combination of these extrema. Figure 1 graphically describes this space, with the personalization dimension increasing from the left to the right and the activity dimension increasing from the bottom to the top of the figure.

*** Figure 1 About Here ***

The point of this diagram is not to locate specific examples precisely in the space, but to help conceptualize the implications of and tradeoffs between the two dimensions in terms of viral feature design. Personalized referrals and notifications, the two examples we empirically evaluate in this paper are denoted in grey boxes, while other examples are denoted in white boxes. Personalized referrals are the most personalized of our examples and also require the most effort by users, while notifications require very little effort beyond normal use of the product and are not personalized but instead broadcast to a wider population. Each of these examples could also ‘move’ in the space depending on the specific instantiation of the viral feature. For instance, notifications that only target a subset of the population, a user’s personal social network or specific individuals (through collaborative filtering for example) would each appear further to the right of the automated notification example noted in the diagram. Examples of personalized referrals include Facebook, Gmail and Plaxo invitations. Examples of automated notifications include Facebook notifications, FourSquare geo-location notifications and the failed Facebook Beacon system which broadcast users’ product purchases to peers. Generalized hypertext embedding, which includes examples such as YouTube, Slide.com and RockYou.com require more effort to post than notifi-

\(^4\) Personalization is a function of both targeting and customization, two attributes of a viral feature that are likely to covary.
Personalized hypertext embedding, such as profile box installations on Facebook, are more personalized than generalized hypertext embedding because they target a user’s personal social network rather than the general population of Internet browsers. Collaborative bookmarking sites like Delicious.com are personalized but also include an element of algorithmic activity. In that sense the user is only partly responsible for activating the feature and so we place collaborative bookmarking to the right of hypertext embedding as it is more personalized. The automated targeted notifications box represents a feature that could exist if notifications were targeted toward specific individuals using collaborative filtering. As we do not know of any features that combine these two processes, we have used a dotted line to mark the box on the figure.

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”

---

5 The potential audience reached by generalized hypertext embedding is subject to efforts of self-selection (since users must choose to visit the location containing the embedded link), popularity of the location, and popularity of the posting individual. Typically, the size of this audience can be many times larger than the typical number of social network peers.
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. Taken together, these arguments lead to the following testable hypotheses:

Hypothesis 1 (H1). Enabling viral product design features increases the likelihood of adoption among peers of current users.

Hypothesis 2 (H2). Viral product design 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.

Hypothesis 3 (H3). Viral product design features that require more activity on the part of the user and are more personalized to recipients generate fewer total viral messages.

3. Empirical Methods

3.1. 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. If uncorrected, these biases can lead researchers to incorrectly attribute observed correlations to the influence of individuals on their peers, resulting in misinterpretations of the treatment effects of viral marketing campaigns or viral product design choices.

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 longi-
tudinal 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 1993). 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 (1967) 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.6 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

---

6 The processes the result in homophily are sometimes also referred to as “selection” effects or “endogenous tie formation.”
assumptions, be identified using instrumental variables based on mean deviations or structural differences (Frank and Strauss 1986, Bramoullé et al 2009, Kelejian and Prucha 1998, Lee 2003, 2006, 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, Snijders et al 2009). 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 Darlouf 2001, Sacredote 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. For example, Tucker (2008) estimates the effect of the adoption of video streaming technologies in the UK on the adoption likelihoods of peers of those in the UK who themselves live in the US. She argues that the World Cup represents an exogenous event that should boost the utility from video streaming for those in the UK, but that this boost should not affect those in the US because the appeal of football matches to the average person living in the UK is higher.
than in the US. However, it could be that those in the US that choose to have friends in the UK are much more likely to be football fans than the average American, demonstrating the difficulty of teasing apart selection (or endogenous tie formation) from influence. 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, although ad hoc methods can convincingly corroborate theories of causal peer influence, it is difficult to establish their robustness with authority. Christakis and Fowler (2007) suggest that the directionality of ties can support causal arguments about peer influence. For example, they argue that because they observe transfers of weight gain over time in the opposite direction of friendship nominations, that obesity is induced by peers. However, it could simply be that popular individuals (whom others nominate as friends more than they nominate others) have higher self esteem and are therefore at lower risk of becoming obese. The point of this review is not to dwell on the weaknesses of these methods (they are all improvements over correlation analysis and important for understanding causation in observational data), but rather to underscore the difficulty of identifying peer effects and to motivate our use of a randomized field experiment to recover estimates of peer influence and social contagion.

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, Leider et al 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 (Angrist, Imbens and Rubin 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). To robustly implement our randomized trial, 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.

3.2. Experimental Design and Procedures

We partnered with a firm that develops commercial applications hosted on the popular social networking site Facebook.com to elicit data on the peer influence effects of enabling viral features using a commercial application built for use on the Facebook platform. Facebook is an ideal environment in which to study peer influence for four reasons. First, experiments on Facebook can capture natural user behavior and do not suffer from a potential loss of external validity that may affect laboratory experiments in which users are removed from their normal daily routines. Second, experiments conducted on Facebook can tap the broad audience of Facebook’s user base which consists of hundreds of millions of individuals that interact on a daily basis and collectively participate in tens of billions of relationships. Third, such experiments can log detailed digital records of users’ online representations and interactions, such as demographics, preferences, social networks, online behavior, and product adoption decisions. Fourth, the Facebook application development environment can be leveraged to control experimental treatment conditions in exacting detail.

The application we studied provides users the opportunity 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

---

7 This is known as the “Stable Unit Treatment Value Assumption” (see Angrist, Imbens and Rubin 1996).
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.8

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, Imbens and Rubin 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 effectiveness of different viral messaging channels, as well as exploration of the mechanisms by which a particular viral channel influenced recipient behavior.

Facebook allows application developers the ability to implement a variety of viral features that send communications to Facebook peers through well-defined channels. The two primary viral features that were implemented in the application and enabled or disabled for different experimental treatment conditions are described below and illustrated using the Flixster Facebook application, an example application that is similar to the one used to conduct the experiment.9

Notifications. 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

8 Facebook allows users to specify privacy settings that may restrict an application’s access to some or part of their profile. This is unlikely to have a significant effect on the study, as it is estimated that less than 2% of Facebook users alter default privacy settings (Gross et al 2005).

9 Due to confidentiality considerations, we do not reveal the application or firm.
peers and displayed in a status bar at the bottom of the peers’ Facebook environment as shown in Figure 2. 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. Because 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 personalized message, they are classified as low personalization (broadcast) in the viral feature space.

***Figure 2 About Here***

*Personal Referrals or Invitations (Invites).* 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 so, they are taken to the application canvas page where they are given the opportunity to install the application. This process is illustrated in Figure 3. 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 personalization than notifications in the viral feature space.

***Figure 3 About Here***

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 treatment group received a version of the application in which both notifications and invites were disabled. In the passive-broadcast treatment group (passive), only notifications were enabled. In the active-personalized treatment group (active), both notifications and invites were enabled. There were no other differences between baseline, passive and active applications. Throughout the experiment, each adopter of the application was randomly assigned to a treatment group according to the proportions displayed in Table 1. The proportion of users assigned to the baseline was chosen in agreement with the ap-
application developer to obtain a population size sufficient to establish a comparative baseline, while limiting potential adverse effects on the overall diffusion of the product that were deemed undesirable by the application developer.10

*** Table 1 About Here ***

Throughout the experiment detailed logs of application user activity, adoption times, viral feature use, peer response, and application user and peer profile data were recorded. Additionally, social network relationships for application adopters and mutual ties between peers of application users were recorded. Our experimental design enabled us to measure the effect of the treatment on the adoption response of peers of treated users as displayed in Figure 4.

*** Figure 4 About Here ***

Recruitment. 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 firm operates as an ad-exchange agency that delivers advertisements through dedicated advertising spaces within a large number of Facebook applications that span a broad range of contexts and user bases. 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 $1.60 per recruit.11 The number of impressions, clicks, and install responses to the recruitment campaign are displayed in Table 2. Summary statistics of the recruited study population are described in § 4. Application diffusion in three randomly selected baseline, passive and active users’ local networks is shown in Figure 5.

---

10 The developer feared too many baseline users could stunt the viral diffusion of the application, which made us even more interested to test whether these features actually caused social contagion. They therefore insisted that the number of baseline users be limited. Limiting baseline users should in no way bias our results as the proportion of baseline users to either treatment group is constant across treatments. Limiting baseline users should only make our estimates more conservative in that analyses comparing a treatment group to baseline will have less power.

11 The cost per recruited user is incredibly low—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.
4. Measurement, Analysis and Results

4.1. 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 73,253 viral messages to their peers, resulting in 992 peer adoptions in direct response to viral messages. Three main observations arise from consideration of the summary statistics of the resultant data displayed in Table 3.

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 across the treatment and 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. For example, 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. Measures of the speed of adoption in a treated user’s local network, as indicated by the time to the first, second, third and fourth adoption events reveal that the treatments increase the rate of adoption in a treated user’s local network. For example, 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 extent to which the effect of the treatment leads to adoption beyond a user’s immediate local network can be measured by the maximal diffusion depth, the maximal distance in the social graph from a treated user to a peer adopter. The average maximal diffusion depth 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. T-tests reveal these differences to be highly significant.

Finally, the extent to which treatment leads to increased application use is measured by users’ average application activity. 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. It is not clear from these cursory examinations alone which explanation is correct or which mechanisms are driving influence and contagion. We therefore turn in the next section to several more sophisticated analyses.

4.2. Effects of Viral Product Design on Peer Influence and Social Contagion

We use a number of different modeling techniques to estimate the peer influence effects of randomly assigned viral features on application adoption and diffusion in the networks of experimental and control users. Our main statistical approach uses discrete-time hazard modeling, which is the standard
Creating Social Contagion through Viral Product Design

A technique for assessing contagion in economics, marketing, and sociology literatures (e.g. Coleman 1967, Van den Bulte and Lilien 2001, Iyengar et al 2008, 2010, Nam et al 2009). This approach typically represents the hazard of adoption of individual $i$ at time $t$ as:

$$P(y_{it} = 1 | y_{i(t-1)} = 0) = F(x_{it} y_{it} , \beta \sum_j w_{ij} y_{jt}),$$

where $y_{it} = 1$ if $i$ has adopted the product by time $t$, and $y_{it} = 0$ otherwise; $F$ is a cumulative distribution function; $x_{it}$ is a vector of variables unrelated to social influence that affect $i$'s adoption decision; $w_{ij}$ is the social exposure of $i$ to peer $j$; and $\gamma$ and $\beta$ are parameters to be estimated.

Discrete time hazards 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. In order to estimate the effect of a randomly assigned application on the adoption of peers, working from the individual outward to the social environment is the only option. Controlled “treatments” of each user’s social environment are too complex and costly to be accomplished reliably in the field 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 (rather than estimating the effects of the ‘outside’ social environment ‘inward’ on the user) 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 therefore 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). In our case, 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. Another important characteristic of the process that produced these data is that the number of adoption events $K_i$ for individual $i$ is a random component of the data generating process and is therefore informative of the distribution of recurrence times (Chang and Wang 1999). 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:

\[ \hat{\lambda}_k(t, X_{ki}) = \hat{\lambda}_{0k}(t)e^{\beta'X_{ki}}, \]

where stratification occurs over the \( K \) adoption events, \( \hat{\lambda}_{0k}(t) \) represents the baseline hazard of the \( k^{th} \) adoption event (i’s \( k^{th} \) 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 \( \beta \) is a vector of unknown parameters to be estimated (Prentice et al. 1981). We assume i’s \( k^{th} \) 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 \( \beta \) using standard maximum likelihood estimation and adjust the covariance matrix to account for non-independence across individuals \( i \) using the following robust covariance matrix:

\[ V = I^{-1}GG'I^{-1} \]

where \( G \) is a matrix of group efficient residuals. Results are presented in Table 4.

*** Table 4 About Here ***

Table 4, 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 Hypothesis 2 – more personalized active features have a greater marginal impact on the rate of peer adoption 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, which records each time stamped viral message sent and any response to it by peers of experimental and control users confirm these results. Table 5 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 in converting new users.

*** Table 5 About Here ***

These results corroborate the main hazard rate results and together confirm Hypotheses 1, 2 and 3: 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.

*** Figure 6 About Here ***

Figures 6a) and 6b) plot the cumulative peer adoptions and the fractions of adopters in the local networks of baseline, passive and active treatment users, while 3d) plots the Kaplan-Meier survival estimates for baseline, passive and active treatments respectively.\(^{12}\) Susceptible peers of users in the passive-broadcast viral treatment group had an approximately seven-fold higher fraction of adopters in their local networks compared to baseline users. Susceptible peers of users in the active-personalized treatment group had over a ten-fold increase in adoption fraction compared to users in the baseline group, and an additional 1.5-fold increase in adoption fraction over peers of users in the passive viral treatment group. These graphs confirm that viral feature design has an economically significant impact on the diffusion of product adoption. Figure 6c) shows that treated users also use the application more than baseline users, which suggests that a positive feedback loop may be driving social contagion. In the next section we focus on the social and economic processes underlying peer influence and contagion and examine evidence that suggests that positive externalities may be accelerating contagion.

4.3. Positive Feedback – Mechanisms Driving Social Contagion

Results of hazard models present strong evidence of social contagion effects caused by viral product design features. However, although the randomized experiment provides unbiased causal esti-

\(^{12}\) Figure 3b) plots the fraction of susceptible peers that adopt the application \(t\) days after they become susceptible in active-personalized, passive-broadcast and baseline treatment and control groups, while Figure 3a) shows the cumulative adoption in each group. To assess the effect of the treatment group on the adoption of application user’s peers through any influence-mediating channel, we identify the time of susceptibility to influence for all peers of buy-in users. To account for fixed-time effects, we look at the adoption response of all susceptible peers \(t\) days after they first became susceptible. We define the adoption fraction as, \(A_j(t)\):

\[
A_j(t) = \frac{\text{Number of susceptible peers that have adopted } t\text{ days after becoming susceptible}}{\text{Number of peers that are still susceptible } t\text{ days after becoming susceptible}}
\]

and we plot the adoption fraction as a function of \(t\) for peers of buy-in users assigned to the baseline, passive, and active viral treatment groups.
mates of the average treatment effects of turning on active and passive viral features, it cannot explain how these features generate social contagion, or what Van den Bulte and Stremersch (2004: 532) call the “nature of the diffusion process.” Several social processes could underlie the dramatic impact of viral features on social contagion in our data. For instance, it could be that the features themselves make the application more interesting and therefore drive application use and peer adoption. Akin to the arguments of those that study which “characteristics” make a product ‘go viral,’ this mechanism would predict that viral features are correlated with both peer adoption and use, but that additional adoption by peers would not necessarily increase the rate of future adoption. On the other hand, network externalities, social pressure, competition and social learning could all generate a positive feedback loop of use and additional peer adoption (Van den Bulte and Stremersch 2004). In this section we examine the details of our rich dataset in more depth to flesh out the underlying social processes that are likely to be driving the social contagion we observe. In particular we examine whether there is a positive feedback loop that accelerates contagion.

We begin by examining the distributional form of the diffusion process over time. Although “one cannot distinguish between contagion and heterogeneity only on the basis of statistical properties of the distributional form,” “the use of such data can improve and refine causal analyses…” (Taibleson 1974: 878) Taibleson continues (1974: 878) “analyses over time and between intervals may help define the precise character of a given causal structure.” We are cautious not to push our data too hard or to ask too much of our longitudinal data given that we have a limited period of observation and relatively few adoption events. However, consistent with prior theory on the distributions of reinforcement or contagion processes (e.g. Feller 1943, Arbous and Kerrich 1951, Coleman 1964, Allison 1980, Holden 1986), the hazard rate of adoption is increasing over adoption events in our data corroborating the results of the randomized controlled experiment. The reinforcement effect seems to increase faster than exponentially for the first several adoption events, then more slowly suggesting that it is approximately constant over time (Allison 1980).

***Table 6 About Here***
Although distributional properties of the reinforcement rate can help characterize a given causal structure, and although the functional form of diffusion processes can help distinguish contagion mechanisms given enough precise data (Van den Bulte and Stremersch 2004), we interpret these results with some caution and instead explore the relationships between covariates in more depth.

We investigate alternate social explanations by examining relationships between application features, activity and use, and application diffusion. As we do not have the benefit of randomization beyond the original application features, we rely on observational analyses to shed light on underlying mechanisms. Table 7 presents results of OLS regression models estimating how these factors correlate with the number of peer adopters, the depth of diffusion and the level of 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, again confirming Hypothesis 2 and providing a robustness check on the original model specification.\textsuperscript{13}

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. Taken together, the distributional properties of the baseline hazards of adoption events and evidence of a strong correlation between the number of adopter friends and application use suggest that a positive feedback loop exists that accelerates contagion – as more of a user’s friends adopt, they use they application more creating positive reinforcement.

Two caveats about this final conclusion are worth noting. First, a natural question arises – if such a positive feedback loop exists, why did the application not spread exponentially ‘infecting’ everyone on Facebook in a short period of time? We speculate that this is because there are a limited number of people who are interested in this product. Though the spreading process is accelerating as adoption events occur, limited interest will always attenuate the spread. Also, this result does not consider retention. As users stop using the application, its diffusion will be curtailed. Second, although evidence of the existence of a

\textsuperscript{13} We also estimated regressions on the percentage of adopters in a treated user’s local network, which yielded qualitatively similar results.
positive feedback loop exists, we do not believe we can robustly distinguish whether it is caused by positive network externalities, social pressure, competition or social learning. Future work is needed to adjudicate between these alternate explanations in social contagion research.

5. Discussion

On February 26, 2010, Facebook announced that as of March 1, 2010 they will no longer deliver broadcast notifications from applications and will discontinue support for the sending of notifications from applications to users. Instead of using the current version of universal notifications, developers will in the future be able to choose from a number of different features to communicate with users, including “counters,” “dashboard news notices,” “add bookmark” features and “email request” features (Warren 2010). Counters will replace notifications but will be specific to each application, rather than being universal, which represents an attempt to increase the personalization of the feature to be more tailored to specific users’ interests. The move to encourage email message communication with those that opt in to such messages is an attempt to move further to the right along the personalization dimension of the space of viral features. Ostensibly, these changes, which are part of a larger set of efforts designed to “help improve the quality of [developers’] communications with users” are geared to “help make interactions with applications more streamlined, clear, and less spammy for users.” (Winters 2010) But they are also examples of attempts by Facebook to engineer and design the features of their platform in ways that optimize user interactions. It is hard to imagine that Facebook is not aware of the marginal effectiveness of broadcast features such as notifications compared to more personalized features such as invites and email. If one considers the social cost of “spammy interactions,” continuous redesign of social interaction features for the purpose of optimizing the user experience is likely a rational, profit maximizing strategy. Optimization from Facebook’s perspective may take place over several different variables and constraints, but improving the virality and use of the applications available to users is likely one of their goals. The more

14 See Winters (2010) “Continuing to Simplify Facebook”:
users use Facebook, the more valuable the platform becomes both in terms of continued growth and in terms of advertising effectiveness. These moves by Facebook are clear examples of viral product design aimed at engineering user interactions to increase sharing, interaction and the virality of products launched on their platform. A cursory survey of the literature on two sided markets reveals how clearly firms profit by promoting the adoption of complementary goods due to network externalities enabled by the complements (Parker and Van Alstyne 2005, Boudreau Forthcoming). Viral product design is a way to enable the uptake of such complementary goods.

The difficulty however is in determining what works and what does not. Causal relationships between policies and outcomes in social processes are inherently difficult to decipher. Numerous statistical challenges prevent clean econometric estimation of the likely effects of changes in product design features and platform policy. But, IT based products, services and platforms provide a natural arena for experimentation. Companies like Amazon, Yahoo and Facebook conduct “A-B” tests of website and user interface design constantly to understand how minor changes affect user behavior and profitability. Viral product design is only made easier by the ability to test and experiment with different viral features, to assess the effects on the magnitude and duration of peer influence and the distributional form of reinforcement effects created by different features and their use. Given the low cost of conducting such experiments, the rapid development and testing of viral design features and the winner take all nature of markets with network externalities (Katz and Shapiro 1985, 1986, Farrell Saloner 1992), we believe that this type of experimentation and testing will only increase in the future and eventually become commonplace in the development of many IT-based products and services and beyond.

6. Limitations and Future Work

There are three important limitations of the study that deserve note. First, as with any study of social contagion that only considers one product or service, the nature of the product is important. Although we seek to measure the effect of incorporating viral features into existing products, there may be some applications to which our results cannot be generalized. We selected an application that is represen-
tative of typical application products developed for the Facebook environment. However, our results may not generalize to an atypical product, such as one that offers limited repeat usability or has no social function. In addition, while the product we studied has many of the characteristics that IT products and services incorporate, whether these results generalize to non-IT related products is an important consideration. For example, running shoes are not typically thought of as IT enabled products and services and if one is considering whether our results apply to the market for running shoes, a first thought might be that the scenarios are too different to make appropriate comparisons. However, we believe that more and more traditional products are becoming IT enabled. For example, the Nike Plus system incorporates digital sensors into running shoes that track movements, record workouts and collect data for those who wish to analyze their own progress. An entire web community as been built around the Nike Plus system where runners can go to share stories, recommend ‘good runs’ and so forth. We can imagine viral features being built into such a system and in fact such features are being built. For example, you can invite your friends to compete with you on the website based on your recorded data. Today, most products are or could become IT enabled and viral features are applicable to most if not all of them.

Second, we have coarse data on friendship ties. Our data only record whether two people are Facebook friends. These are undirected, not weighted and there is little indication of the strength or nature of the ties. Future work should seek to understand how different types of ties and the strength of ties mediate the effects of viral product features on peer outcomes. Even in Facebook, finer data on relationships could be constructed. For example, tagged images establish collocation and a given date and time (as well as the fact that two individuals were willing to be photographed together). Such data could be used to discern real friendships from cursory acquaintances.

Third, we do not observe the content of communication, nor do we observe channels of communication outside of Facebook. Previous research demonstrates the importance of observing and analyzing content in network data (Aral and Van Alstyne 2009). These omissions should only serve to make our results more conservative, but they also limit our ability to explore the social mechanisms and processes that underlie contagion.
Creating Social Contagion through Viral Product Design

Our results highlight the importance of several lines of future research beyond those we have already mentioned. For example, understanding which individuals are influential and which individuals are susceptible in social contagions that arise due to the use of viral product design features may prove to be important. Certainly experimental methods can make great strides in helping to resolve recent debates about whether the influentials hypothesis holds in non-observational data. In addition, an important element that has been left in the background in our work is the importance of network structure. As our focus is on the relational elements of social contagion (how one person may influence another) we do not foreground network structure. Important questions around the disproportional influence of hubs and especially the globally diffusion properties of viral feature influence propagation are as yet unexplored. Finally, and critically, we only examine influence and contagion in product adoption decisions. A critical element in the propagation and diffusion of new products is their sustained use. We encourage more work on how and when viral features and influence in general sustain interest in a product such that customer churn does not overwhelm new adoptions.

7. 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. 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.

References


Creating Social Contagion through Viral Product Design


Goldenberg J, Lowengart O, & Shapira D (2010) integrating the social network to diffusion model and evaluation of the value of hubs in the adoption process.


Creating Social Contagion through Viral Product Design


Creating Social Contagion through Viral Product Design


Winters (2010) “Continuing to Simplify Facebook”:
### Tables and Figures

#### Table 1. Stratification Across Treatment Groups

<table>
<thead>
<tr>
<th></th>
<th>Baseline Control</th>
<th>Passive-broadcast Treatment</th>
<th>Active-personalized Treatment</th>
</tr>
</thead>
<tbody>
<tr>
<td>5%</td>
<td>47.5%</td>
<td>47.5%</td>
<td></td>
</tr>
</tbody>
</table>

#### Table 2. Recruitment Statistics Describing the Initial Advertising Campaign

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

#### Table 3. Summary Statistics and Mean Comparisons of Active, Passive and Baseline Users

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean (SD)</td>
<td>Mean (SD)</td>
<td>Mean (SD)</td>
<td>t-statistic (SE)</td>
<td>t-statistic (SE)</td>
<td>t-statistic (SE)</td>
</tr>
<tr>
<td>Age</td>
<td>31.51 (13.80)</td>
<td>30.81 (13.31)</td>
<td>29.94 (13.27)</td>
<td>.46 (13.35)</td>
<td>1.03 (13.31)</td>
<td>1.45 (13.24)</td>
</tr>
<tr>
<td>Gender (1 = Male)</td>
<td>.25 (.44)</td>
<td>.33 (.47)</td>
<td>.32 (.47)</td>
<td>-1.57 (.47)</td>
<td>-4.2 (.46)</td>
<td>.4 (.47)</td>
</tr>
<tr>
<td>Degree†</td>
<td>171.79 (223.88)</td>
<td>170.25 (278.64)</td>
<td>166.97 (248.77)</td>
<td>.09 (275.13)</td>
<td>.32 (247.15)</td>
<td>.55 (263.82)</td>
</tr>
<tr>
<td>Number of Facebook Wall Posts</td>
<td>40.52 (79.89)</td>
<td>36.45 (223.88)</td>
<td>37.07 (246.76)</td>
<td>.46 (93.11)</td>
<td>.15 (238.20)</td>
<td>.09 (188.31)</td>
</tr>
<tr>
<td>User’s Local Network</td>
<td>.01 (.12)</td>
<td>.07 (.35)</td>
<td>.10 (.44)</td>
<td>-2.84 (.34)</td>
<td>-3.6 (.43)</td>
<td>-3.64 (.40)</td>
</tr>
<tr>
<td>Number of Adopters in User’s Local Network</td>
<td>.02 (.02)</td>
<td>.09 (.01)</td>
<td>.15 (.01)</td>
<td>-1.92 (.01)</td>
<td>-2.35 (.01)</td>
<td>-2.83 (.01)</td>
</tr>
<tr>
<td>Percentage of Adopters</td>
<td>.01 (.01)</td>
<td>.09 (.1)</td>
<td>.15 (.1)</td>
<td>-2.53 (.21)</td>
<td>-3.01 (.24)</td>
<td>-1.98 (.23)</td>
</tr>
<tr>
<td>Maximum Diffusion Depth</td>
<td>4.4 (.11)</td>
<td>4.77 (.22)</td>
<td>3.17 (.24)</td>
<td>1.27 (.21)</td>
<td>2.04 (.24)</td>
<td>2.45 (.23)</td>
</tr>
<tr>
<td>Time to 1st Adopter</td>
<td>9.4 (.71)</td>
<td>4.77 (.84)</td>
<td>3.17 (.62)</td>
<td>1.27 (.80)</td>
<td>2.04 (.67)</td>
<td>2.45 (.70)</td>
</tr>
<tr>
<td>Time to 2nd Adopter</td>
<td>5.23 (8.17)</td>
<td>5.29 (6.97)</td>
<td>3.04 (6.72)</td>
<td>1.27 (8.07)</td>
<td>2.45 (6.77)</td>
<td>2.45 (7.30)</td>
</tr>
<tr>
<td>Time to 3rd Adopter</td>
<td>6 (8.07)</td>
<td>6 (.87)</td>
<td>3.04 (5.25)</td>
<td>1.27 (5.07)</td>
<td>2.45 (5.25)</td>
<td>2.45 (6.33)</td>
</tr>
<tr>
<td>Time to 4th Adopter</td>
<td>6 (8.07)</td>
<td>6 (.87)</td>
<td>3.04 (5.25)</td>
<td>1.27 (5.07)</td>
<td>2.45 (5.25)</td>
<td>2.45 (6.33)</td>
</tr>
<tr>
<td>Application Activity</td>
<td>3.17 (4.59)</td>
<td>4.17 (2.24)</td>
<td>4.56 (8.98)</td>
<td>-2.54 (7.08)</td>
<td>-2.89 (8.73)</td>
<td>-2.20 (8.16)</td>
</tr>
</tbody>
</table>

Notes: ***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.; .01, p = .94, N.S.
Table 4: Variance-Corrected Proportional Hazards of Contagion in Networks of Baseline, Passive and Active Treatment Groups

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Viral State = Passive</strong></td>
<td><strong>Viral State = Active</strong></td>
<td><strong>Application Activity</strong></td>
<td><strong>Notifications</strong></td>
</tr>
<tr>
<td>Hazard Ratio (SE)</td>
<td>Hazard Ratio (SE)</td>
<td>Hazard Ratio (SE)</td>
<td>Hazard Ratio (SE)</td>
</tr>
<tr>
<td>3.46*** (1.18)</td>
<td>3.35*** (1.15)</td>
<td>2.50** (.86)</td>
<td>2.51** (.86)</td>
</tr>
<tr>
<td>4.44*** (1.64)</td>
<td>4.21*** (1.56)</td>
<td>3.33*** (1.24)</td>
<td>3.31*** (1.24)</td>
</tr>
<tr>
<td>1.02*** (.004)</td>
<td>1.02*** (.003)</td>
<td>1.02*** (.002)</td>
<td>1.02*** (.003)</td>
</tr>
<tr>
<td>Log Likelihood</td>
<td>-4694.359</td>
<td>-4631.795</td>
<td>-4544.845</td>
</tr>
<tr>
<td>(X^2) (d.f)</td>
<td>19.34*** (2)</td>
<td>57.41*** (3)</td>
<td>298.78*** (4)</td>
</tr>
<tr>
<td>Observations</td>
<td>3929</td>
<td>3929</td>
<td>3929</td>
</tr>
</tbody>
</table>

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

Table 5: Click Stream Analysis of Responses to Viral Messages and Adoption

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Invitations</strong></td>
<td><strong>Notifications</strong></td>
<td><strong>Adoption Rate (Marginal Impact)</strong></td>
</tr>
<tr>
<td>Messages Sent</td>
<td>Adoptions via Click Through Installation</td>
<td></td>
</tr>
<tr>
<td>160</td>
<td>16</td>
<td>.10</td>
</tr>
<tr>
<td>69980</td>
<td>666</td>
<td>.001</td>
</tr>
</tbody>
</table>

Table 6: Baseline Hazards Over \(k\) Events \(\hat{\lambda}_0(k = 1...6)\)

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Mean (SD)</strong></td>
<td><strong>Min</strong></td>
<td><strong>Max</strong></td>
<td><strong>N</strong></td>
</tr>
<tr>
<td>(\hat{\lambda}_{01})</td>
<td>.0002 (.0001)</td>
<td>.0001</td>
<td>.001</td>
</tr>
<tr>
<td>(\hat{\lambda}_{02})</td>
<td>.002 (.001)</td>
<td>.001</td>
<td>.013</td>
</tr>
<tr>
<td>(\hat{\lambda}_{03})</td>
<td>.015 (.024)</td>
<td>.005</td>
<td>.14</td>
</tr>
<tr>
<td>(\hat{\lambda}_{04})</td>
<td>.034 (.010)</td>
<td>.021</td>
<td>.054</td>
</tr>
<tr>
<td>(\hat{\lambda}_{05})</td>
<td>.046 (.008)</td>
<td>.037</td>
<td>.067</td>
</tr>
<tr>
<td>(\hat{\lambda}_{06})</td>
<td>.099 (.044)</td>
<td>.053</td>
<td>.14</td>
</tr>
</tbody>
</table>
### Table 7. Drivers of Application Diffusion

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Beta (SE)</td>
<td>Beta (SE)</td>
<td>Beta (SE)</td>
<td>Beta (SE)</td>
<td>Beta (SE)</td>
<td>Beta (SE)</td>
<td>Beta (SE)</td>
<td>Beta (SE)</td>
<td>Beta (SE)</td>
<td>Beta (SE)</td>
</tr>
<tr>
<td>Viral State = Passive</td>
<td>.078** (.031)</td>
<td>.084** (.033)</td>
<td>.020 (.059)</td>
<td>.045** (.0178)</td>
<td>.048*** (.019)</td>
<td>.020 (.018)</td>
<td>.129* (.074)</td>
<td>.112 (.079)</td>
<td>.062 (.076)</td>
<td>.037 (.074)</td>
</tr>
<tr>
<td>Viral State = Active</td>
<td>.119*** (.031)</td>
<td>.131*** (.032)</td>
<td>.059* (.030)</td>
<td>.057*** (.018)</td>
<td>.063*** (.019)</td>
<td>.033* (.018)</td>
<td>.190*** (.074)</td>
<td>.171** (.079)</td>
<td>.091 (.076)</td>
<td>.006 (.074)</td>
</tr>
<tr>
<td>Degree</td>
<td>.0001*** (.00002)</td>
<td>.0001** (.0003)</td>
<td>.0001** (.00002)</td>
<td>.0001*** (.00001)</td>
<td>.00004** (.00002)</td>
<td>.00007** (.00001)</td>
<td>.00007** (.00001)</td>
<td>.00002** (.00001)</td>
<td>.00002** (.00001)</td>
<td>.00002** (.00001)</td>
</tr>
<tr>
<td>Facebook Activity</td>
<td>.019*** (.006)</td>
<td>.06 (.006)</td>
<td>.061*** (.005)</td>
<td>.010*** (.0004)</td>
<td>.005*** (.0002)</td>
<td>-.003 (.006)</td>
<td>.054*** (.016)</td>
<td>.042*** (.015)</td>
<td>.026* (.014)</td>
<td>--</td>
</tr>
<tr>
<td>Application Activity</td>
<td>.019*** (.006)</td>
<td>.06 (.006)</td>
<td>.061*** (.005)</td>
<td>.010*** (.0004)</td>
<td>.005*** (.0002)</td>
<td>-.003 (.006)</td>
<td>.054*** (.016)</td>
<td>.042*** (.015)</td>
<td>.026* (.014)</td>
<td>--</td>
</tr>
<tr>
<td>Notifications</td>
<td>.035*** (.010)</td>
<td>.035*** (.010)</td>
<td>.035*** (.010)</td>
<td>.035*** (.010)</td>
<td>.035*** (.010)</td>
<td>.035*** (.010)</td>
<td>.035*** (.010)</td>
<td>.035*** (.010)</td>
<td>.035*** (.010)</td>
<td>.035*** (.010)</td>
</tr>
<tr>
<td>Invites</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Number of Adopters</td>
<td>.607*** (.030)</td>
<td>.607*** (.030)</td>
<td>.607*** (.030)</td>
<td>.607*** (.030)</td>
<td>.607*** (.030)</td>
<td>.607*** (.030)</td>
<td>.607*** (.030)</td>
<td>.607*** (.030)</td>
<td>.607*** (.030)</td>
<td>.607*** (.030)</td>
</tr>
<tr>
<td>F Value (d.f.)</td>
<td>12.20*** (3)</td>
<td>11.18*** (4)</td>
<td>157.94*** (7)</td>
<td>9.36*** (3)</td>
<td>10.11*** (4)</td>
<td>85.13*** (7)</td>
<td>3.51*** (3)</td>
<td>4.87*** (4)</td>
<td>83.54*** (5)</td>
<td>128.92*** (7)</td>
</tr>
<tr>
<td>R²</td>
<td>.006</td>
<td>.007</td>
<td>.16</td>
<td>.004</td>
<td>.007</td>
<td>.09</td>
<td>.002</td>
<td>.003</td>
<td>.07</td>
<td>.14</td>
</tr>
<tr>
<td>Observations</td>
<td>8910</td>
<td>5766</td>
<td>5766</td>
<td>6310</td>
<td>5766</td>
<td>5766</td>
<td>6310</td>
<td>5766</td>
<td>5766</td>
<td>5766</td>
</tr>
</tbody>
</table>

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

Figure 1. The Viral Product Feature Space.
Figure 2. A Graphical Example of the Interface and Process of the Notifications Feature.

Users can invite their friends to adopt the application and join their social network on the application itself.

Figure 3. A Graphical Example of the Interface and Process of the Personal Referrals or Invitations Feature.

Figure 4. Graphical Representation of the Experimental Comparison
Figure 5. Three Representative Local Networks of a) Baseline, b) Passive and c) Active Users. The Initial experimental adopter (ego) is colored dark purple, peer adopters (friends of ego who adopted) are colored red, peer of peer adopters (friends of friends of ego who adopted) are colored orange. Non-adopters are white.
Figure 6. 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.
Figure 7. Baseline Hazards ($\lambda_{0k}$) for $k = 1\ldots6$ fitted to an Exponential and a Power Function