MANAGEMENT SCIENCE
Vol. 57, No. 9, September 2011, pp. 1623–1639
ISSN 0025-1909 | EISSN 1526-5501 | 11 | 5709 | 1623

http://dx.doi.org/10.1287/mnsc.1110.1421

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

Sinan Aral, Dylan Walker
Information, Operations, and Management Sciences, Stern School of Business, New York University, New York, New York 10012 {sinan@stern.nyu.edu, dwalker@stern.nyu.edu}

We examine how firms can create word-of-mouth peer influence and social contagion by designing viral features into their products and marketing campaigns. To econometrically identify the effectiveness of different viral features in creating social contagion, we designed and conducted a randomized field experiment involving the 1.4 million friends of 9,687 experimental users on Facebook.com. We find that viral features generate econometrically identifiable peer influence and social contagion effects. More surprisingly, we find that passive-broadcast viral features generate a 246% increase in peer influence and social contagion, whereas adding active-personalized viral features generate only an additional 98% increase. Although active-personalized viral messages are more effective in encouraging adoption per message and are correlated with more user engagement and sustained product use, passive-broadcast messaging is used more often, generating more total peer adoption in the network. Our work provides a model for how randomized trials can identify peer influence in social networks.

Key words: peer influence; social contagion; social networks; viral marketing; viral product design; information systems; randomized experiment
History: Received February 28, 2010; accepted March 31, 2011, by Pradeep Chintagunta and Preyas Desai, special issue editors. Published online in *Articles in Advance* August 4, 2011.

1. Introduction

It is widely believed that social contagion and word-of-mouth (WOM) "buzz" about products drive product adoption and sales, and firms increasingly rely on "network" and "viral" marketing strategies (Hill et al. 2006, Manchanda et al. 2008, Nam et al. 2010). Yet, whereas most current work has focused on viral marketing campaigns for existing products, less attention has been paid to whether (and how) firms can design products that are themselves more likely to go viral. The effectiveness of such *viral product design* strategies have yet to be examined or causally estimated. We therefore conducted a large-scale randomized field experiment to test the effectiveness of different viral product design features in creating peer influence and social contagion in new product diffusion.

Viral product design—the process of explicitly engineering products so they are more likely to be shared among peers—has existed at least since the first chain letter was sent in 1888. Today, products regularly use information technology (IT)-enabled features like automated broadcast notifications and personalized invitations to spread product awareness. Yet, although viral features have become more sophisticated and a central part of the design of products and marketing campaigns, there is almost no empirical evidence on the effectiveness of such features in generating social contagion and product adoption. We therefore investigate two basic questions: Can firms add viral features to products so they are more likely to be shared among peers? If so, which viral features are most effective in inducing WOM and peer-to-peer influence in product adoption?

Unfortunately, evaluating the effects of viral product design features is difficult because peer effects and WOM are typically endogenous (Manski 1993; Godes and Mayzlin 2004, 2009; Hartmann et al. 2008; Aral et al. 2009; Aral 2011). We therefore designed and conducted a randomized field experiment testing the effectiveness of two of the most widely used viral product features—active-personalized referrals and passive-broadcast notifications—in creating peer influence and social contagion among the 1.4 million friends of 9,687 experimental users of Facebook.com. The experiment uses a customized commercial Facebook application to observe user behavior, communications traffic, and the peer influence effects of randomly enabled viral messaging features on application diffusion and use in the local networks of experimental and control population users. By

enabling and disabling viral features among randomly selected users, we were able to obtain relatively unbiased causal estimates of the impact of viral features on the adoption rates of peers in the local networks of adopters. Using detailed clickstream data on users' online behaviors, we also explored whether positive network externalities generated by additional peer adopters inspired further product adoption and sustained product use.

WOM is generally considered to be more effective at promoting product contagion when it is personalized and active. Surprisingly, we find that designing products with passive-broadcast viral messaging capabilities generates a 246% increase in local peer influence and social contagion, whereas adding active-personalized viral messaging capabilities generates only an additional 98% increase. Although active-personalized messaging is more effective in encouraging adoption per message and is correlated with more user engagement and sustained product use, it is used less often and therefore generates less total peer adoption in the network. Overall, we find that viral product design features generate econometrically identifiable peer influence and social contagion effects and provide a model for how randomized trials can identify peer influence in networks.

2. Viral Product Design

Since the early work of Katz and Lazersfled (1955) there has been great interest in how WOM drives consumer demand, public opinion, and product diffusion (Brown and Reingen 1987, Godes and Mayzlin 2004, Aral et al. 2009) and how firms can create broad, systematic propagation of WOM through consumer populations (Phelps et al. 2004, Mayzlin 2006, Dellarocas 2006, Godes and Mayzlin 2009). Many campaigns target "influential" individuals who are likely to propagate organic WOM most broadly (Katz and Lazersfeld 1955, Watts and Dodds 2007, Goldenberg et al. 2009), using referral programs to create incentives for them to spread the word (Biyalogorsky et al. 2001). Others use observational evidence on viral campaigns to inform viral branching models of WOM diffusion (Van der Lans et al. 2010). However, to this point, studies of viral product design have remained conspicuously absent from the literature on viral marketing.

Viral product design involves incorporating specific characteristics and features into a product's design to generate peer-to-peer influence that encourages adoption. A product's *viral characteristics* are fundamentally about its content and the psychological effects content can have on a user's desire to share the product with peers (Stephen and Berger 2009, Berger and Heath 2005, Heath et al. 2001). A product's *viral features*, on the other hand, concern how the product is shared—how features enable and constrain a product's use in relation to other consumers. Viral features may enable communication, generate automated notifications of users' activities, facilitate personalized invitations, or enable hypertext embedding of the product on publicly available websites and weblogs. Two of the most widely used viral product features are personalized referrals and automated broadcast notifications:

Personalized Referrals. Personalized referral features allow users to select their friends or contacts from a list and invite them to adopt the product or service, with the option of attaching a personalized message to the invitation. Social networking websites enable users to "invite their friends" to join the service through personalized referrals. When users send Web-based e-mail messages, for example, from Gmail, an automated, pop-up hyperlink enables them to invite recipients to join the service.

Automated Broadcast Notifications. Automated broadcast notifications are passively triggered by normal user activity. When a user engages the product in a certain way (e.g., sends a message, updates his or her status), those actions are broadcast as notifications to the user's list of contacts. Notifications build awareness among friends of new activities or products a user is adopting or engaging with, and can encourage those friends to eventually adopt the product themselves. For example, social networking websites typically notify friends automatically when a user adopts a new application or achieves some application milestone.

Referrals are more personalized and targeted than broadcast notifications. Users actively select a subset of their social network to receive them (targeting) and can include personal messages in the referral (personalization). WOM is generally considered more effective at promoting product contagion when it is personalized and active. When individuals choose to share information about products and services with their friends, they tend to activate their strong-tie relationships (Frenzen and Nakamoto 1993, Aral and Van Alstyne 2011). Strong ties exhibit greater homophily (Jackson 2008), greater pressure for conformity (Coleman 1988), and deeper knowledge about one another. We tend to trust information from close "trusted" sources more and to respond more often to them because of reciprocity (Emerson 1962). In addition, the personalization of messages makes them more effective, especially in online environments in which we are bombarded with irrelevant information (Tam and Ho 2005, Tucker 2010).

For these reasons, one might suspect that personalized referrals are more effective than broadcast notifications. But, although each personalized referral may

be more persuasive (more effective per message), the pervasiveness of automated broadcast messages that do not require additional time and energy on the part of the user may lead to greater overall peer adoption. The effort required by the user to actively select and invite peers to adopt the product may curtail widespread use of the personalized referral and so limit its effectiveness in encouraging broad adoption. The relative overall effectiveness of these viral features is therefore ultimately an empirical question.

3. Experimental Design and Procedures

Evaluating the effects of viral product design features on social contagion is difficult because peer effects and WOM are typically endogenous (Manski 1993, Hartmann et al. 2008, Van den Bulte and Lilien 2001, Godes and Mayzlin 2004, Van den Bulte and Iyengar 2011). Several approaches for identifying peer effects have been proposed, including peer effects models and extended spatial autoregressive models (e.g., Oestreicher-Singer and Sundararajan 2008, Trusov et al. 2009, Bramoulle et al. 2009), actor-oriented models (e.g., Snijders et al. 2006), instrumental variable methods based on natural experiments (e.g., Sacredote 2001, Tucker 2008), dynamic matched sample estimation (Aral et al. 2009), structural models (e.g., Ghose and Han 2010), and ad hoc approaches (Christakis and Fowler 2007). However, randomized trials are considered to be one of the most effective ways to obtain unbiased estimates of causal peer effects (Duflo et al. 2008, Hartmann et al. 2008).

We therefore partnered with a firm that develops commercial applications hosted on the popular social networking website Facebook.com and collected experimental data on the peer influence effects of enabling viral features on the diffusion one of their applications. This application is free to adopt and provides users the opportunity to share information and opinions about movies, actors, directors, and the film industry in general. We designed multiple experimental versions of the application in which *personalized invitations* and *broadcast notifications* were enabled or disabled, and randomly assigned adopting users to various experimental and control conditions. As users adopted the application, each was randomly assigned to one of the two treatment conditions or the baseline control condition. The application collected personal attributes and preferences from users' Facebook profiles, as well as data on their social networks and the personal attributes and preferences of their network neighbors.[1]

The experiment enabled experimental group users to use passive-broadcast and active-personalized viral messaging capabilities to exchange messages with their network neighbors, while disabling those 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 clickstream data on recipient responses to viral messages and their subsequent adoption and use of the application. When an individual adopted the application as a result of peer influence, their treatment status was also randomized to ensure that the stable unit treatment value assumption held. This facilitated analysis of the relative effectiveness of different viral messaging channels in generating peer adoption and network propagation. Randomization also enabled exploration of the mechanisms by which a particular viral channel influenced recipient behavior. Two primary viral features were examined:

Automated Broadcast Notifications (*Notifications*). When enabled, notifications were generated automatically when an application user performed certain actions within the application, such as declaring a favorite movie or writing a movie review. When notifications were generated, they were distributed to a random subset of an application user's peers and displayed in a status bar at the bottom of the peers' Facebook environment. When a peer clicked on the notification, they were taken to an application canvas page where they were given the option to install the application. These notifications required no effort beyond normal application use, making their engagement relatively costless to the user. Because they were randomly distributed to a Facebook user's peers and were not accompanied by a personalized message, they also exhibited low personalization.

Personalized Referrals or Invitations (*Invites*). When enabled, invites allowed application users to send their Facebook peers personalized invitations to install the application. Peers received the invitation in their Facebook inbox and could click on a referral link contained within the invitation. If they did so, they were taken to the application canvas page where they were given the opportunity to install the application. Each invite required a conscious and deliberate action from the user beyond typical application use, requiring more effort (activity) than notifications. Because invites were targeted to specific peers and allowed the inclusion of a personalized message, they also exhibited greater personalization.

[1] 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, because it is estimated that less than 2% of Facebook users alter default privacy settings (Gross et al. 2005).

Table 1 Stratification Across Treatment Groups

Baseline control (%)	Passive-broadcast treatment (%)	Active-personalized treatment (%)
5	47.5	47.5

The experimental design consisted of three treatment groups into which users were randomly assigned: *baseline*, *passive-broadcast*, and *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 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.[2]

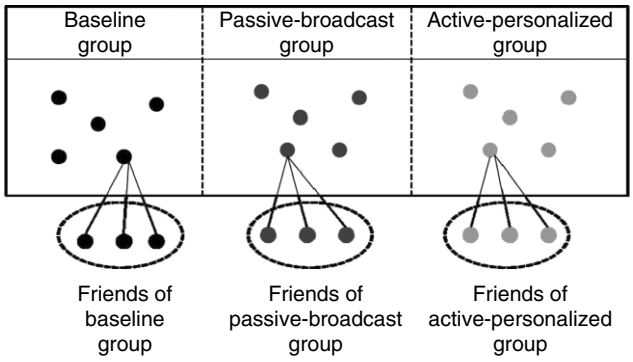
Detailed logs of application user activity, adoption times, viral feature use, peer response, and application user and peer profile data were recorded, as were social network relationships for application adopters and mutual ties between peers of application users. Our experimental design allowed us to measure the effect of each of the viral features on the adoption response of peers, as displayed in Figure 1. We also measured adopters' use by recording activity logs that detailed each time a user took an action on the application.

3.1. Recruitment

At the launch of the experiment, we designed an advertising campaign in collaboration with a second Facebook advertising firm to recruit a representative population of Facebook users. Advertisements were displayed to users through advertising space within Facebook and within existing Facebook applications. The campaign was conducted in three waves throughout the duration of the experiment and cost a total of $6,000 to recruit 9,687 usable experimental subjects,

[2] The developer feared too many baseline users could stunt the viral diffusion of the application and therefore insisted that the number of baseline users be limited. Limiting baseline users should not bias results as the proportion of baseline users to either treatment group is constant across treatments and should only make our estimates more conservative in that analyses comparing a treatment group to the baseline group will have less power.

Figure 1 Graphical Representation of the Experimental Comparison



or 62¢ per recruit. The number of impressions, clicks, and installation responses are displayed in Table 2. Summary statistics of the recruited study population are described in §4. Comparisons to published demographic statistics indicate the sample is indeed representative of typical Facebook users (see the appendix). The application was also publically listed in Facebook's Application Directory and so was available to anyone on Facebook.

4. Empirical Methods

4.1. Data and Descriptive Statistics

The experiment was conducted over a 44-day period, during which 9,687 initial users adopted the application, with 405 users randomly assigned to the baseline control group, 4,600 users randomly assigned to the passive-broadcast treatment group, and 4,682 users randomly assigned to the active-personalized treatment group. Users in these groups collectively had 1.4 million distinct peers in their local social networks and sent a total of 70,140 viral messages to their peers, resulting in 992 peer adoptions—682 of which were 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 in terms of their age, gender, network degree (number of Facebook friends), and level of Facebook activity (number of

Table 2 Recruitment Statistics Describing the Initial Advertising Campaign

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

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

	1 Baseline ($N = 405$)	2 Passive ($N = 4{,}600$)	3 Active ($N = 4{,}682$)	4 *t*-statistic (B-P)	5 *t*-statistic (B-A)	6 *t*-statistic (P-A)
	Mean (SD)	Mean (SD)	Mean (SD)	*t*-statistic (SE)	*t*-statistic (SE)	*t*-statistic (SE)
Age	31.51 (13.80)	30.81 (13.31)	29.94 (13.27)	0.46 (13.35)	1.03 (13.31)	1.45 (13.24)
Gender (1 = male)	0.25 (0.44)	0.33 (0.47)	0.32 (0.47)	−1.57 (0.47)	−1.42 (0.46)	0.40 (0.47)
Degree[a]	171.79 (223.88)	170.25 (278.64)	166.97 (248.77)	0.09 (275.13)	0.32 (247.15)	0.55 (263.82)
Number of Facebook wall posts	40.52 (79.89)	36.45 (94.16)	37.07 (246.76)	0.46 (93.11)	0.15 (238.20)	−0.09 (188.31)
Number of adopters in user's local network	0.01 (0.12)	0.07 (0.35)	0.10 (0.44)	−2.84*** (0.34)	−3.60*** (0.43)	−3.64*** (0.40)
Percentage of adopters in user's local network	0.02 (0.002)	0.09 (0.01)	0.15 (0.01)	−1.92* (0.01)	−2.35** (0.01)	−2.83*** (0.01)
Maximum diffusion depth	0.01 (0.11)	0.04 (0.22)	0.05 (0.24)	−2.53* (0.21)	−3.01*** (0.24)	−1.98*** (0.23)
Time to 1st adopter	9.40 (9.71)	4.77 (8.04)	3.17 (6.72)	1.27 (8.07)	2.04** (6.77)	2.45*** (7.30)
Time to 2nd adopter	—	5.23 (8.17)	4.43 (6.97)	—	—	0.58 (7.45)
Time to 3rd adopter	—	5.29 (8.07)	3.04 (5.25)	—	—	1.08 (6.33)
Time to 4th adopter	—	6 (5.83)	1.17 (1.12)	—	—	2.84*** (3.58)
Application use	3.17 (4.59)	4.17 (7.24)	4.56 (8.98)	−2.54** (7.08)	−2.89*** (8.73)	−2.20* (8.16)

Notes. This table reports means and standard deviations for demographic variables, peer adoption statistics, and Facebook and application activity statistics of baseline (column 1), passive (column 2), and active (column 3) control and treatment group users, as well as results of *t*-tests of mean differences between baseline and passive users (column 4), baseline and active users (column 5) and passive and active users (column 6). Variables reported include *Age*: self-reported age on Facebook; *Gender* (1 = male): self-reported gender on Facebook; *Degree*: number of Facebook friends; *Number of Facebook wall posts*: count of the number of "wall posts" posted to an individual's Facebook profile recorded at the beginning of the study; *Percentage of adopters in user's local network*: the percentage of an individual's Facebook friends who adopted the application calculated at the end of the observation period; *Time to 1st, 2nd, 3rd, 4th adopters*: the time in days to the first, second, third, and fourth adopters in the user's friend network; *Application use*: a continuous measure of application calls from a user's account to the application server indicating the number of actions taken on the application.

[a]Kolmogorov–Smirnov tests of degree distribution differences: baseline-passive, 0.04, $p = 0.80$, not significant (NS); baseline-active, 0.04, $p = 0.79$, NS; passive-active, 0.01, $p = 0.94$, NS.

*$p < 0.01$; **$p < 0.05$; ***$p < 0.001$.

Facebook wall posts), confirming the integrity of the randomization procedure.

Second, whereas 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. The number of peer adopters in a user's local network is approximately 7 times greater for users in the passive-broadcast treatment group and 10 times greater for users in the active-personalized group. Similarly, compared to the baseline group, the percentage of adopters in a user's local network is approximately 450% higher for users in the passive-broadcast group and 750% higher for users in the active-personalized group. 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 increased the rate of adoption in a treated user's local network. The time to the first peer adoption is approximately 200% shorter for users in the passive-broadcast treatment group and approximately 300% shorter for users in the active-personalized group. The extent to which the treatment leads to adoption beyond a user's immediate local network can be measured by the maximal diffusion depth—the maximum network distance from a treated user to any peer adopter in a linked chain of adoptions. The average maximal diffusion depth is approximately

360% greater for the passive-broadcast treatment group and 450% greater for the active-personalized treatment group compared to baseline users. T-tests reveal these differences are highly significant.

Finally, the extent to which each treatment leads to increased application use is measured by users' average application activity. Average application activity is approximately 130% higher in the passive-broadcast treatment group and 140% higher in the active-personalized treatment group. We go beyond these initial statistics and explore more formal models of peer influence and social contagion in the next several sections.

4.2. Model Specification

Our main statistical approach uses hazard modeling, which is the standard technique for assessing contagion in economics, marketing, and sociology (e.g., Van den Bulte and Lilien 2001, Iyengar et al. 2011, Nam et al. 2010). This approach typically represents the hazard of adoption for individual i at time t as a function of individual characteristics and social influence:

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

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

Our circumstances, however, required a slightly different approach. We are interested in estimating the treatment effects of randomly assigned viral features on the adoption of peers of experimental and control users, rather than the effects of users' social environments on their own adoption decisions. Controlling users' entire social environments is typically too complex and costly to be accomplished reliably in the field. We therefore adopted an "inside-out" strategy that estimated 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 (see the appendix).

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 in which multiple peer adoptions can occur for the same user over time. Failure times are correlated within users' local networks, violating the assumption of independence of failure times required in traditional survival analysis. Network effects may also make contagion a function of prior adoptions in a local network. We therefore employed a variance-corrected stratified proportional hazards approach that accounts for the lack of independence among the multiple clustered failure times in the data and allows the baseline hazards to vary by adoption event to account for the possibility that adoption hazards vary across stages of a diffusion process. We estimated the following model:

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

in which stratification occurs over the K adoption events, $\lambda_{0k}(t)$ represents the baseline hazard of the kth adoption event (i's kth friend adopting), X_{ki} represents a vector of covariates affecting the adoption of i's neighbors (including i's viral treatment status (active, passive, or baseline), a measure of i's level of activity on the application (application activity), peer notifications sent (notifications), and invites sent (invites)), and β is a vector of unknown parameters to be estimated. If t_{ik} is the adoption time for the kth adoption in i's network, adoption times are sequential such that $t_{ik} \geq t_{ik-1}$. Adoption data are time stamped in minutes and seconds, ensuring no two events happen at the same time. The conditional risk set at time t for event k therefore consists of all subjects under observation at time t who have experienced a $k-1$ adoption event. We estimated β using standard maximum likelihood estimation and adjusted the covariance matrix to account for nonindependence across individuals i using the following robust covariance matrix (where G is a matrix of group efficient residuals):

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

4.3. Robustness

We conducted numerous additional analyses to ensure that our results are robust to selection bias, contamination, information leakage, and multiple alternative specifications of the contagion model. First, selection bias could occur when a user chooses to adopt the experimental application through the recruitment campaign or when they adopt in response to a viral message. We took steps to measure and mitigate both possibilities. The recruitment campaign was designed to reach a representative audience of Facebook users. The demographics of our study population are comparable to those of the broader Facebook population, and published Facebook demographics fall within one standard deviation of study population sample means. In addition, application users that adopt because of peer influence may themselves be subject to selection effects and may therefore be fundamentally different than application users that adopted via initial recruitment. We eliminate these

sources of selection bias by only considering initially recruited users in the randomized treatment group to which they were assigned.

Second, in randomized trials in network environments, users assigned to different treatment groups may not strictly be isolated from one another. This raises the concern that information leakage through indirect network pathways may contaminate the results of the study. It is important to note that in traditional studies, whether or not the network is observed, relationships that create leakage effects may still exist between treatment and control populations. One benefit of our design is that we systematically observe how individuals in the study are connected, enabling us to measure and prevent leakage. Even when making conservative assumptions about the decay of information flow and quality over network distance, we estimate potential leakage effects to be small.[3] However, to ensure robustness, we controlled for leakage by (a) evaluating peer adoption only in the local networks of recruited users and (b) right censoring peers when they gained more than one adopter friend.

Finally, we tested several different functional forms and specifications of the contagion model. Detailed evidence showing the robustness of our results to selection bias, contamination, leakage, and alternative model specifications are provided in the appendix.

5. Results

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

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. Users of the passive-broadcast application experienced a 246% increase in the rate of application adoption by peers compared to the baseline group, whereas adding active-personalized viral messaging capabilities generated only an additional 98% increase (active-personalized users experienced a 344% increase over the baseline group). 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. 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%, whereas one additional notification increases the rate of peer adoption by only 2% on average, confirming that more personalized active features have a greater marginal impact on the rate of peer adoption per message than passive broadcast features.

The clickstream data, which record each time-stamped viral message and any response to it by peers, corroborate these results. Table 5 displays the number of invitations and notifications sent, the responses to those messages that resulted in click-through

[3] Results of these analyses are available from the authors.

Table 4 Variance-Corrected Stratified Proportional Hazards of Contagion in Networks of Baseline, Passive, and Active Treatment Groups

	1 Hazard ratio (SE)	2 Hazard ratio (SE)	3 Hazard ratio (SE)	4 Hazard ratio (SE)
Viral state = passive	3.46***	3.35***	2.50**	2.51**
	(1.18)	(1.15)	(0.86)	(0.86)
Viral state = active	4.44***	4.21***	3.33***	3.31***
	(1.64)	(1.56)	(1.24)	(1.24)
Application use		1.02***	1.02***	1.02***
		(0.004)	(0.003)	(0.003)
Notifications			1.02***	1.02***
			(0.002)	(0.002)
Invites				1.06**
				(0.028)
Log likelihood	−4,694.359	−4,631.795	−4,544.845	−4,542.577
χ^2 (df)	19.34***	57.41***	298.78***	307.47***
	(2)	(3)	(4)	(5)
Observations	3,929	3,929	3,929	3,929

Notes. This table reports parameter estimates and standard errors from the variance-corrected stratified proportional hazards model specified in §4.2 with robust standard errors clustered around users' local network neighborhoods. Variables reported include *Viral state = passive* (a dummy variable denoting passive viral application users), *Viral state = active* (a dummy variable denoting active viral application users), *Application use* (a continuous measure of application calls from a user's account to the application server indicating the number of actions taken on the application), *Notifications* (integer count of the number of notifications sent), *Invites* (integer count of the number of invites sent).

**$p < 0.05$; ***$p < 0.001$.

Table 5 Clickstream Analysis of Responses to Viral Messages and Adoption

	1 Messages sent	2 Adoptions via click through installation	3 Adoption rate (marginal impact)
Invitations	160	16	0.10
Notifications	69,980	666	0.01

installations of the application, and the resultant adoption rate per message. Invitations are the least used feature, but the most effective per message in creating peer influence and social contagion. Notifications, which require the least effort and are sent automatically to randomly selected peers, generate more messages, but are less effective per message in converting new users.

These results together confirm the main findings of the study: viral product design features do in fact generate econometrically identifiable peer influence and social contagion effects. Features that require more activity on the part of the user and are more personalized to recipients create greater marginal increases in the likelihood of adoption per message, but also generate fewer messages, resulting in less total peer adoption in the network.

Figure 2, (a) and (b), plots the cumulative peer adoptions and the fractions of adopters, respectively, in the local networks of baseline, passive, and active treatment users. To assess the effect of the treatment on the adoption of application users' peers through any influence-mediating channel, in this figure we identify the time of susceptibility to influence for all peers of initial adopters. To account for fixed-time effects, we look at the adoption response of all susceptible peers t days after they first became susceptible. Figure 2(d) plots the Kaplan–Meier survival estimates for baseline, passive, and active treatments. Susceptible peers of users in the passive-broadcast viral treatment group had an approximately sevenfold 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 10-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 2 **Plots of (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**

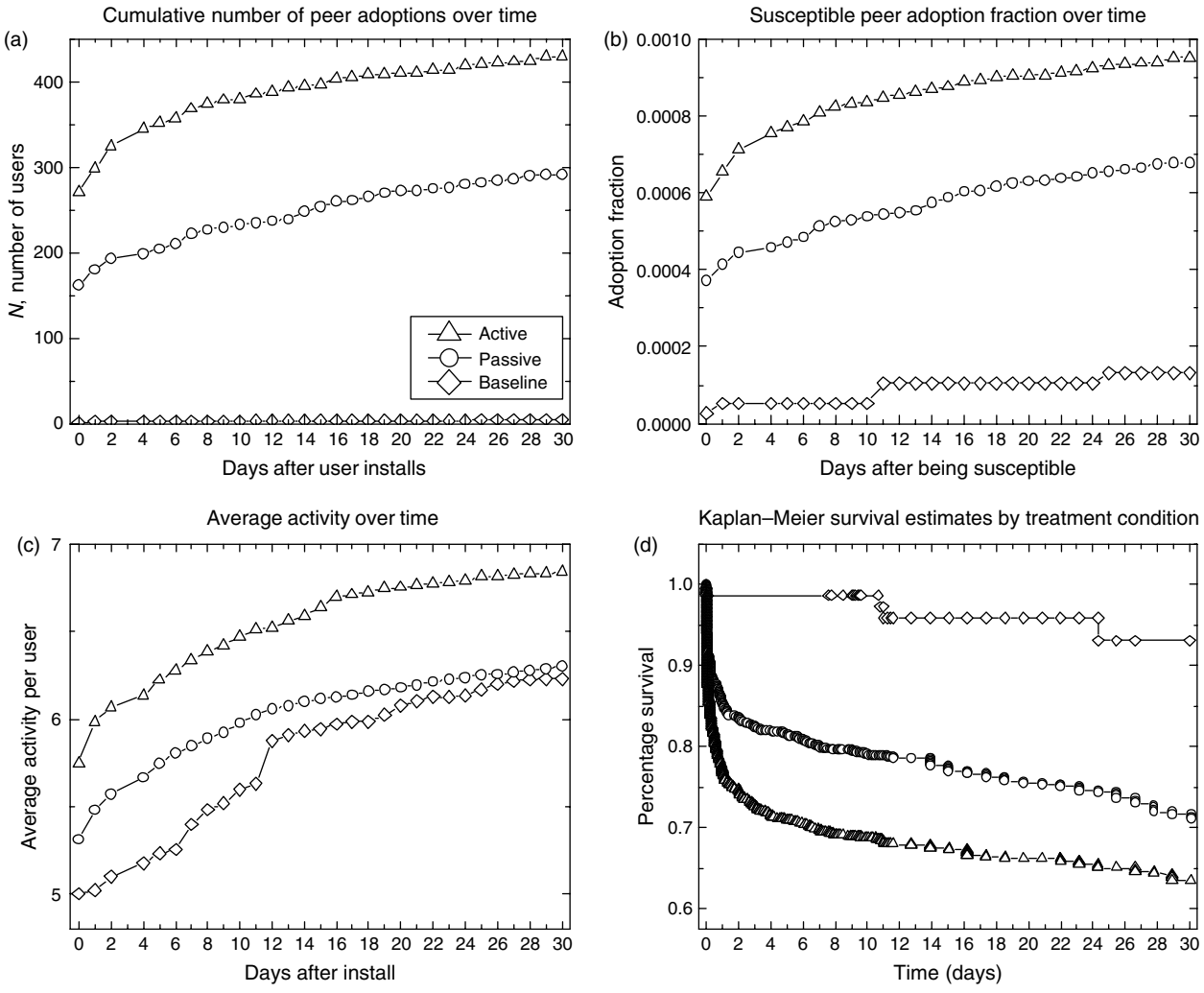
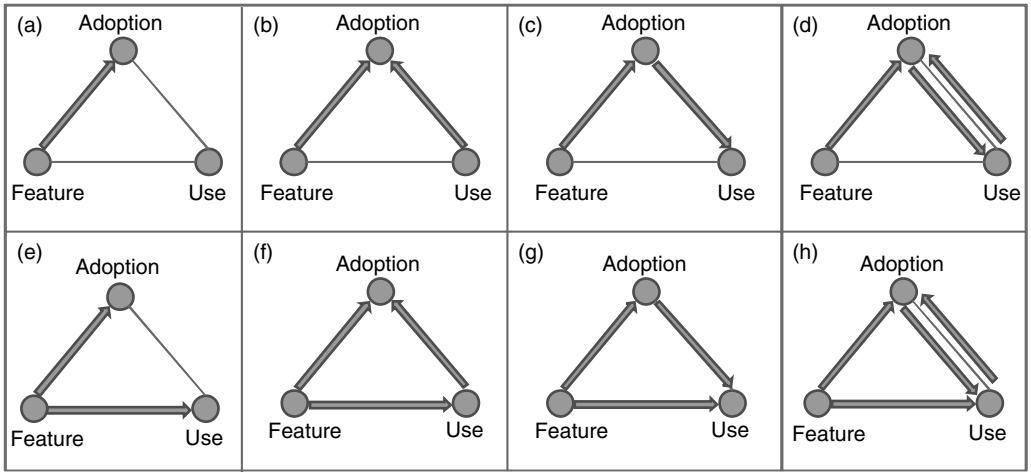


Figure 3 Possible Causal Relationships Between the Existence of Application Features, Peer Adoption, and Application Use



Note. Arrows indicate causal direction.

5.2. Mechanisms Driving Social Contagion

Several social mechanisms could explain how viral features create product contagion. An unexpected result from the experiment enabled us to investigate these mechanisms. Interestingly, treated users not only had more peer adopters, but also used the application more than control group users (see Figure 2(c)). As Figure 2(c) shows, active-personalized users used the application more than passive-broadcast users, who in turn used the application more than baseline users. This result is surprising because users were randomly assigned to different applications, and the versions were identical from users' points of view apart from the invitation option included in the active-personalized application. Understanding why use differed across treatment groups (despite randomization) provides insight into how viral features create contagion and sustained product use.

Viral feature inclusion, application use, and peer adoption are correlated, as shown in Table 8, Model 1, and Figure 2, (a) and (c). The randomized trial confirms that viral features cause peer adoption. We also know that because features are randomized and not controlled by the user, no other covariate can drive the existence of features. Given these constraints, we depict the remaining possible causal relationships between feature inclusion, application use, and peer adoption in Figure 3 and evaluate the possible explanations for these relationships.

Cases (a) and (b) in Figure 3 cannot explain the observed correlation between viral feature inclusion and peer adoption. Specifically, they are inconsistent with the discrepancy in application use between users in different treatment groups (Figure 2(c) and Table 7).

Cases (c) and (d) in Figure 3 represent a network externalities mechanism (Van den Bulte and Stremersch 2004) in which peer adoption drives increased application use by the original adopter. The number of peer adopters a user has is positively associated with their own sustained use of the application even after controlling for their treatment status, degree, and overall Facebook activity (Table 7, Model 3). In addition, users of active-personalized and passive-broadcast applications exhibit more use (Table 7, Model 1), again controlling for observable differences in users' overall Facebook activity (Table 7, Model 2). These results are consistent with the existence of positive network externalities—as more of their peers adopted the application, users were more engaged and used the application more.[4] Although passive-broadcast features are associated with more product use than the baseline early on, this association disappears over time (see Figure 2(c)). That active-personalized features are associated with sustained product use over time and passive-broadcast features are not may suggest a direct network effect from interacting with specific peers—those that are personally invited to the application by their friend. Whether these effects are stronger with invited peers or not, the network effects explanation is broadly supported by the evidence and seems plausible.

However, there could be alternative explanations of these results: (1) that users derive utility from use of the viral features, (2) that users derive utility from the mere existence of the viral features, (3) that users are more satisfied when they receive the viral features they expected, and (4) that other omitted variables create a spurious correlation between

[4] Network externalities can take many forms, some of which are mediated by viral features that increase awareness of the application. Contact the authors for more details and analysis of different types of network externalities in this setting.

Table 6 Baseline Hazards Over k Events λ_{0k} ($k = 1 \ldots 6$)

	1 Mean (SD)	2 Min	3 Max	4 *N*
λ_{01}	0.0002 (0.0001)	0.0001	0.001	523
λ_{02}	0.002 (0.001)	0.001	0.013	128
λ_{03}	0.015 (0.024)	0.005	0.14	42
λ_{04}	0.034 (0.010)	0.021	0.054	20
λ_{05}	0.046 (0.008)	0.037	0.067	15
λ_{06}	0.099 (0.044)	0.053	0.14	7

Notes. This table reports means, standard deviations, and minimum and maximum values for baseline hazard lambda parameters of the kth adoption events in users' networks, $k = 1 \ldots 6$.

Table 7 Correlates of Application Use

	1 Application use	2 Application use	3 Application use	4 Application use
	Beta (SE)	Beta (SE)	Beta (SE)	Beta (SE)
Viral state = passive	0.129* (0.074)	0.112 (0.079)	0.062 (0.076)	−0.037 (0.074)
Viral state = active	0.190*** (0.074)	0.171** (0.079)	0.091 (0.076)	−0.006 (0.074)
Degree	−0.0001 (0.0001)	−0.0001 (0.0001)	−0.0002** (0.0001)	−0.0002** (0.0001)
Facebook activity		0.054*** (0.016)	0.042*** (0.015)	0.026* (0.014)
Notifications				0.022*** (0.001)
Invites				0.055** (0.024)
Number of adopters			0.607*** (0.030)	0.360*** (0.031)
F value	3.51***	4.87***	83.54***	128.92***
(df)	(3)	(4)	(5)	(7)
R^2	0.002	0.003	0.07	0.14
Observations	6,310	5,766	5,766	5,766

Notes. This table reports ordinary least squares parameter estimates for a linear estimating equation regressing *application use* (defined in Table 3) on the variables listed, including *Facebook activity* (the normalized sum of integer counts of the number of wall posts, activities, affiliations, groups, interests, pages, notes, favorite books, movies, music, and TV shows, calculated at the beginning of the study). For all other variable definitions, see Tables 3 and 4. Models are estimated with robust standard errors clustered around users' local network neighborhoods.

*$p < 0.10$; **$p < 0.05$; ***$p < 0.001$.

peer adoption and use. We evaluated each of these explanations in turn.

First, cases (e) and (f) in Figure 3 represent a demand effect explanation in which the correlation between features and application use is explained by an increased utility from the viral features. In one variant of the demand effects explanation, use of the features themselves make the application more interesting and therefore simultaneously drives application use and peer adoption, creating a spurious correlation between the two. If use of the invite and notification features was correlated with both application use and peer adoption, and if peer adoption itself was not driving use, the correlation between the number of peer adopters and application use should disappear once we control for the use of invites and notifications. However, when we hold constant application use associated with both notifications and invites, there is still a strong positive relationship between the number of peer adopters and application use (Table 7, Model 3). This suggests that viral feature use does not fully explain the correlation between peer adoption and use.

Second, it could still be that the mere existence of features rather than their use increases users' utility. But, the data do not support this explanation either. When the viral states are entered into the regression they significantly predict application activity in the expected directions and magnitudes (Table 7, Model 1). When the number of peer adopters is controlled, those relationships disappear completely (Table 7, Models 3 and 4), indicating that the viral state of the application alone—or the utility from simply being able to notify or invite friends—does not explain application use.

Third, it could be that users are more satisfied when they receive the viral features that they expected, creating a demand effect from the expected utility of having viral features enabled. We performed an additional analysis to address this alternative explanation. Application use by peer adopters is a reasonable proxy for their satisfaction with the product—the extent to which their expectations regarding the product conform to the product they actually received upon adoption. We therefore examined the application use of peers that adopted through response to a viral message and divided them into two groups: those that received (through random assignment) a version of the product with the ability to send viral messages of the type they received from their influencing peer and those who received a version of the product without the ability to send viral messages of the type they received from their influencing peer. T-tests show that use of the application by those who received the features they would have expected to receive and those that were "disappointed" (so to speak) by not receiving the features they would have expected to receive show no significant differences in application use ($t = 0.9054$; SD = 8.0377). It is therefore unlikely that adoption and

use are explained by expected utility from the existence of viral features.

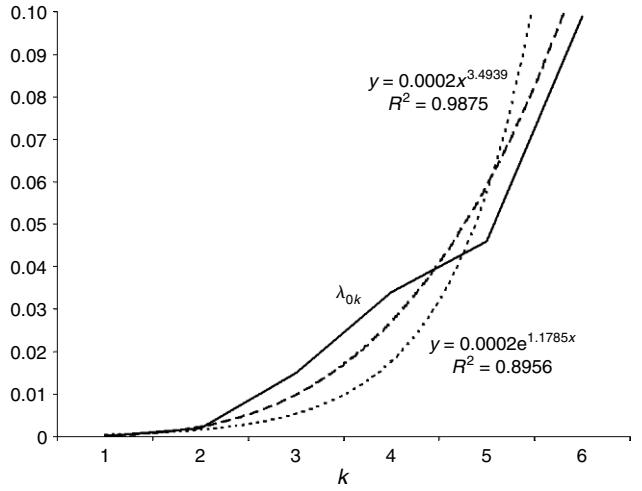
Fourth, other omitted variables (such as unobserved user heterogeneity) could create a spurious correlation between peer adoption and use. However, our data also rule out this possibility. Because feature inclusion is randomized, the distribution of any unobserved covariates must be the same across treatment groups, and so omitted variables cannot produce the discrepancy in peer adoption and application use across treatment groups shown in Figure 2, (a) and (c). There could be an unobserved covariate that must first be activated by the existence of a feature to drive peer adoption and application use; however, that too is unsupported by the evidence. Because we observe correlation between adoption and use beyond that which is explained by use of the invite and notifications features themselves (see Table 7), it seems unlikely that a user characteristic that simultaneously drives peer adoption and use would be activated by a viral feature that users do not use.

The only remaining explanations are depicted in Figure 3, (g) and (h), which captures both demand effects (features drive use) and network externalities (peer adoption drives use). Given our analyses, network externalities are at least in part responsible for driving application use. But, this is a conservative interpretation of the evidence. Because we have ruled out most of the plausible demand effect explanations (that features drive use in one of several ways we have considered), it is likely that network effects are entirely responsible for the increased application use we observe among treated users.

Another piece of evidence corroborating network effects is that the hazard rate of adoption is increasing over adoption events, implying a reinforcement effect of prior adoptions on the likelihood of future adoption (Van den Bulte and Stremersch 2004). The hazard rate of adoption increases faster than exponentially for the first several adoption events, then more slowly, suggesting that reinforcement is approximately constant over peer adoptions (see Table 6 and Figure 4). Although we interpret this piece of evidence with caution because "one cannot distinguish between contagion and heterogeneity only on the basis of statistical properties of the distributional form" (Taibleson 1974, p. 878), the fact that the hazard rate of adoption is increasing with each subsequent adoption event k is consistent with a reinforcement effect of prior adoptions on future adoption.

Finally, Table 8 presents correlates of application diffusion that corroborate results of the randomized trial. Models 4–6 confirm that peers of initial adopters also use the application because diffusion depth depends on peers' (and peers of peers) application use. Active-personalized and passive-broadcast treatments significantly increase average diffusion depth, and these effects are again explained by application use and the viral features themselves (Model 6). Results in Model 3 also corroborate hazard model estimates, confirming that invitations are on average three times more effective per message in inspiring peer adoption than notifications. Taken together, evidence of a strong correlation between the number of adopter friends and application use and the distributional properties of the baseline hazards of adoption events suggest that network externalities accelerate contagion. As more of a user's friends adopt, they use the application more, creating a positive feedback loop.

Figure 4 **Baseline Hazards (λ_{0k}) for $k = 1 \ldots 6$ Fitted to an Exponential and a Power Function**



6. Discussion

These results have broad implications for managers attempting to promote viral product diffusion and for theories of social contagion, opinion leadership, and viral product design. We discuss several of the broad implications of our findings for managers and future research on peer influence, social contagion, and viral marketing.

First, our estimates imply that viral product design may be more effective in encouraging new product adoption than traditional marketing strategies. Notifications and invites, which produced 1% and 10% conversion rates (CRs), respectively, vastly outperform the ad campaign used in our recruitment phase on Facebook, which produced a 0.01% conversion rate. Because Facebook currently has the largest market share of display advertising on the Web, these comparisons reflect the relative performance of viral product design and the lion's share of Web-based display advertising (Tucker 2010). Notifications and

Table 8 Correlates of Application Diffusion

	1 Number of adopters	2 Number of adopters	3 Number of adopters	4 Diffusion depth	5 Diffusion depth	6 Diffusion depth
	Beta (SE)	Beta (SE)	Beta (SE)	Beta (SE)	Beta (SE)	Beta (SE)
Viral state = passive	0.078**	0.084**	0.020	0.045**	0.048***	0.020
	(0.031)	(0.033)	(0.059)	(0.0178)	(0.019)	(0.018)
Viral state = active	0.119***	0.131***	0.059*	0.057***	0.063***	0.033*
	(0.031)	(0.032)	(0.030)	(0.018)	(0.019)	(0.018)
Degree	0.0001***	0.0001**	0.0001**	0.0001***	0.00004**	0.00003**
	(0.00002)	(0.00003)	(0.00002)	(0.00001)	(0.00002)	(0.00001)
Facebook activity		0.019***	0.006		0.013***	0.007**
		(0.006)	(0.006)		(0.004)	(0.004)
Application use			0.061***			0.021***
			(0.005)			(0.003)
Notifications			0.010***			0.005***
			(0.0004)			(0.0002)
Invites			0.035***			−0.003
			(0.010)			(0.006)
F Value	12.20***	11.18***	157.94***	9.36***	10.11***	85.13***
(df)	(3)	(4)	(7)	(3)	(4)	(7)
R^2	0.006	0.007	0.16	0.004	0.007	0.09
Observations	8,910	5,766	5,766	6,310	5,766	5,766

Notes. This table reports ordinary least squares parameter estimates for linear estimating equations regressing the *number of adopters* (defined in Table 3) and *diffusion depth* (the maximum network distance from a treated user to any peer adopter in a linked chain of adoptions) on the variables listed. See Table 7 for all additional notes and variable definitions.

invites also outperform the CR in paid search advertising, which has been estimated at 0.02% (Ghose and Yang 2009). Although conversion rates are typically significantly smaller than click-through rates (CTRs), the CRs for notifications and invites even outperform published statistics on CTRs for traditional banner advertising (outside of Facebook) and e-mail marketing campaigns. The 1% CR on notifications outperforms the CTRs for traditional banner advertising (which range from 0.10%–0.20% in publicly available statistics), and invitations are 10 times as effective at generating conversions as traditional banner ads are at generating click-throughs. Compared to e-mail campaign CTRs (which range from 2% to 6% in publicly available statistics), invitations are again more effective at a 10% conversion rate.[5] These comparisons show viral channels to be more effective at generating higher response rates than traditional digital advertising channels.

We also asked the directors of the firm with whom we partnered about their feature implementation and customer acquisition costs and learned that invites can be implemented for a total cost under $600. Because implementing viral features incurs a low one-time fixed cost, and the expected return is proportional to the increase in adopters the feature generates, viral product design may be a more cost-effective strategy than increasing spending on traditional digital advertising (which incurs costs proportional to impressions or clicks). It may be, however, that the success of viral product design efforts depends on traditional advertising to the extent that an initial base of users is needed to implement viral marketing. It is also important to consider the social cost of viral messages. Bombarding users with messages from peers may reduce the overall quality of the user experience. Future work should estimate the costs of viral product design more comprehensively and consider the implications of both marginal revenue and marginal cost on optimal product design.

Second, given that active-personalized features are more marginally effective but less globally effective than passive-broadcast features, a natural question is how managers can optimize the effectiveness of these viral features. Because the main limitation of active-personalized features is that high effort costs curtail their use, one solution may be to couple active-personalized features with referral incentives that encourage their use (Biyalogorsky et al. 2001). Optimally designed incentive strategies could encourage users to generate more personalized referrals and to target and personalize viral messages more effectively (Aral et al. 2011). It may also be possible to improve the low marginal effectiveness of passive-broadcast

[5] Click-through rates on banner ads have declined from 0.33% to 0.19% from 2004 through 2008 (Riley 2009); DoubleClick (2009) reports that in 2008, the average CTR in the United States was 0.10% for banner ads. For e-mail campaigns, the estimated CTR in 2008 and 2009 remained stable at 5.9% (Epsilon 2009), whereas MailerMailer (2010) reports the average CTR on e-mail campaigns at 2.80%, and WebMarketCentral (2007) reports a 2%–3% CTR.

features by automatically targeting and personalizing broadcast messages algorithmically. If there is a social cost to viral messages, product and platform developers could seek to limit impersonal messages in adaptive ways that are tied to the effectiveness of the messages themselves.

Third, the combination of viral features and network externalities seems to drive a positive feedback loop in which product use drives peer adoption, and peer adoption in turn drives product use. If so, managers should seek to enable this feedback loop by designing both viral features and strong direct and indirect network externalities into their products. Maximizing engagement and minimizing churn may be obvious goals, but the combined effects of engagement and contagion on product diffusion are less obvious. Interactions between network externalities, sustained use, and customer churn may change over a product's life cycle and may vary across products. More work on the relationships between social contagion, sustained use, and customer churn over products and product life cycles will help clarify when viral marketing is most effective.

7. Conclusion

We conducted a large-scale randomized experiment testing the effectiveness of viral product design features in creating social contagion. We found that viral product design has econometrically identifiable impacts on peer influence and social contagion in new product diffusion. Results of our randomized trial suggest that designing viral features into products can increase social contagion by up to 400%. Surprisingly, designing products with passive-broadcast viral messaging capabilities generates more total peer influence and social contagion than adding active-personalized viral messaging capabilities. Although active-personalized messaging is more effective in encouraging adoption per message and is correlated with more user engagement and sustained product use, it is used less often, and therefore generates less total peer adoption in the network. Data on the distributional form of the diffusion process and on product use are consistent with the existence of positive network externalities that reinforce peer adoption and create a virtuous cycle of engagement and contagion.

Understanding optimal viral product design strategies, taking into account factors such as sustained product use, network externalities, social and economic costs, incentives, and the marginal effectiveness of different viral features could enable firms to optimally create and manage social contagion. The difficulty, however, is in determining what works and what does not. Numerous statistical challenges prevent clean causal estimation of the relationships between interventions and outcomes and the likely effects of changes in product design and platform policy. Fortunately, IT-based products and platforms provide natural vehicles for randomized experimentation. Given the low cost of conducting experiments, the rapid development and testing of viral design features, and the winner-take-all nature of markets with network externalities, this type of experimentation is likely to increase in the future and eventually to become commonplace in the development of many products and platforms. Our work sheds light on how viral products can be designed to generate social contagion and offers a template for using randomized trials to identify peer influence in networks.

Acknowledgments

For valuable comments, the authors are grateful to Erik Brynjolfsson, Haim Mendelson, Markus Mobius, Lev Muchnik, Foster Provost, Sean Taylor, Catherine Tucker, Hal Varian, and seminar participants at the National Bureau of Economic Research (NBER), the Workshop on Information in Networks (WIN), the Workshop on Information Systems Economics (WISE), the Sunbelt Social Networks Conference, the International Conference on Network Science (NetSci), the International Conference on Information Systems (ICIS), the Microsoft Research–Cambridge Workshop on Networks and Economics, the Winter Conference on Business Intelligence, Google, Yahoo!, Facebook, Northeastern University, Northwestern University, New York University, Massachusetts Institute of Technology, the University of Minnesota, the University of Michigan, Georgia Tech, and the University of Pittsburgh. Sinan Aral gratefully acknowledges generous funding and support from Microsoft Research (Faculty Fellowship 2010), IBM (Faculty Fellowship 2009), and the National Science Foundation (Career Award 0953832).

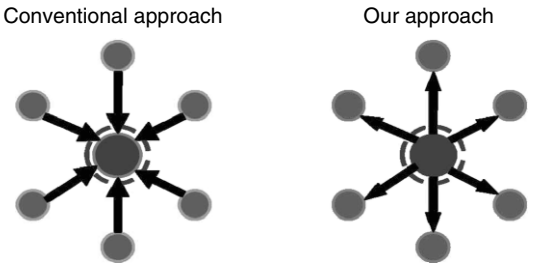
Appendix

Inside-Out Design

Randomized trials are traditionally used to estimate the effect of a treatment on the treated. To study the effect of viral feature incorporation on product adoption outcomes, we instead examined the effect of treatment on the peers of treated application users. The difference in these approaches is illustrated in Figure A.1. Arrows indicate the potential flow of influence that the experiment is designed to detect. The solid blue circle in the center represents the treated user, and the red outlines indicate measurements of treatment effects. In social network environments, a conventional approach is infeasible because it is difficult to comprehensively control the network environments of each user in the study population. It is feasible, however, to treat a user and observe the effect of treatment on the outcomes of their peers.

The strength of our approach lies in its ability to capture effects of any form of influence-mediating communication channels between the treated user and her peers, including effects that arise through influence-mediating communication channels beyond those that can be explicitly recorded. For example, treated users could communicate with and

Figure A.1 Inside-Out Experimental Design



influence their peers through offline interactions such as face-to-face communications or telephone conversations, as well as through unrecorded online communications such as e-mail or external chat conversations. Because we measure the response of peers regardless of how they may or may not have been influenced by treated users, we are able to capture the effect of unrecorded influence-mediating communications on peer adoption.

Preventing Selection Effects

Selection effects could occur when a user chooses to adopt the experimental application through the recruitment campaign or when they adopt in response to a viral message. We took steps to mitigate and measure both possibilities. The recruitment campaign was designed to reach a representative audience of Facebook users, and advertisements were displayed to users through advertising space within Facebook and within existing Facebook applications. Establishing to what extent the recruited population is representative of the general Facebook population is somewhat challenging because Facebook does not officially publish demographic statistics of their user base. However, through the use of a recently released social targeting advertisement service provided by Facebook, it is possible to obtain some official demographic statistics. Age and gender demographics sampled through this application programming interface and published online by istrategylabs.com are comparable to the same demographic ranges for recruited study population users in Figure A.2. Though our sample has a slightly higher percentage of women than the Facebook population, and users in our sample have a slightly higher average degree (150 compared with 130 in Facebook statistics), the demographics of our study population are comparable to those of the broader Facebook population, and the published Facebook demographics fall within one standard deviation of study population sample means.

In addition to issues of selection surrounding the population of recruited users, application users that adopt because of peer influence may be subject to selection effects and may be fundamentally different from application users that adopted via initial recruitment. It could be that users who use the viral features and peers of users who use viral features are systematically different from randomly selected Facebook users. We avoid these sources of selection bias in our analyses by only considering initially recruited users in the randomized treatment group to which they were assigned. Peers of recruited users only contribute to local network peer adoption of originally recruited users and are not themselves used as test subjects.

Preventing Leakage and Contamination

In randomized trials in network environments, users assigned to different treatment groups may not be strictly isolated from one another. This raises the concern that information leakage through indirect network pathways may contaminate the results of the study. It is important to note that in traditional studies, whether or not the network is measured, relationships may still exist between treatment and control populations that create leakage effects. One benefit of our design is that we systematically observe how individuals in the study are connected, enabling us to measure and prevent leakage.

Several factors reduce the likelihood that leakage is affecting our results. First, because treatment assignment is randomized, any leakage will be uncorrelated with treatment assignment and cannot account for the observed differences in responses to treatments. Although it is possible that leakage will on average provide some common information to peers of treated users uniformly across the treatment designations, this effect should only serve to make our estimates across treatment groups more conservative, because leakage should reduce differences between control and treatment groups. Second, information flows between individuals in a network typically decay rapidly with network distance (Wu et al. 2004, Aral et al. 2007). Although all users may be connected through long friendship paths, leakage will diminish over successive hops in each path.

Nonetheless, leakage effects could downward bias our estimates of treatment effects toward zero, and we therefore

Figure A.2 Comparison of Sample and Population Demographic Characteristics

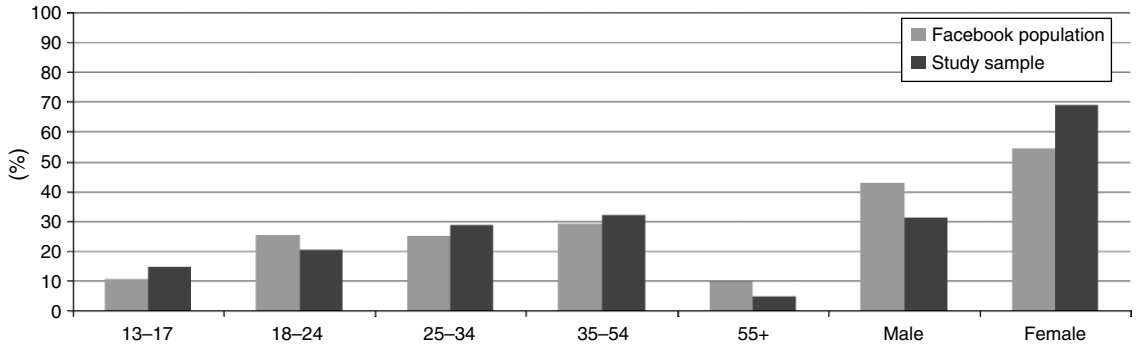
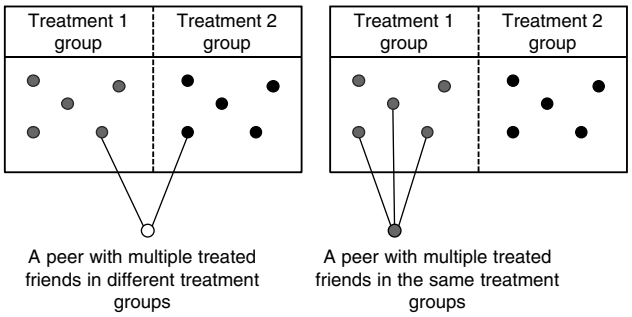


Figure A.3 Preventing Contamination and Leakage



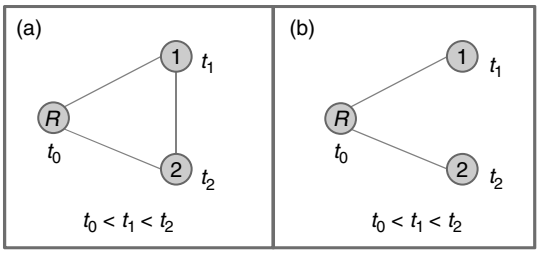
take several steps to prevent leakage. First, in hazard rate models, we examine only peers of initially recruited adopters. In addition to avoiding potential selection issues mentioned above, this also excludes individuals (and their potential adopter peers) that adopt in chains within a local neighborhood, lessening leakage effects. Such individuals are likely to share more and shorter indirect paths with existing adopters than a randomly chosen peer, as a consequence of clustering and mutuality (Newman 2003).

Second, we account for users with multiple treated peers (of similar and/or different treatments). Existence of peers of multiple treated application users leads to two potential complications. First, users may be peers of multiple treated users from different treatment groups, making it impossible to link their treatment effects to a single treatment. Second, peers of multiple treated users that belong to the same treatment group are clearly classified as peers of either baseline, active-personalized, or passive-broadcast users; however, measurements of their response may be incorrectly estimated as a consequence of being subject to influence from multiple treated friends. A peer with multiple treated friends in a given viral treatment group may exhibit an adoption outcome or time to adoption that is systematically different from those of peers with only one treated friend. These two scenarios are displayed in Figure A.3.

The nature of treatment randomization does not allow us to simultaneously guarantee that all treated friends of a peer will receive the same treatment. Consequently we treat peers with multiple treated friends as contaminated as soon as they become so and exclude them from our analysis. This procedure could underestimate the effect of clusters of adoption on the time to adoption or number of adopters in a local network neighborhood; however, if this is the case, it will do so in a manner that is the same for all treatment types. Furthermore, because treatment groups are randomized, there can be no systematic correlation between the type of treatment received by a user and that received by her subsequent adopter peer.

The procedure that we adopt for designating a peer as contaminated is detailed in Figure A.4. The initially recruited adopter, labeled R, adopts at time t_0. Two peers of user R, labeled 1 and 2, adopt at subsequent times t_1 and t_2, respectively. In panel (a), for times $t > t_1$, peer 2 has multiple treated peers (R and 1) that may have been assigned different treatments. Peer 2 is therefore considered contaminated for times $t > t_1$. In panel (b), a similar situation occurs, but no link exists between peers 1

Figure A.4 Procedure for Designating Contaminated Peers



and 2, and consequently neither user is considered contaminated. In our analysis, when a peer is designated as contaminated, she is removed from the hazard rate model for subsequent time periods. This procedure appropriately retains the maximal empirical support for hazard rate estimation and parameterizes our ignorance of what might happen subsequent to a user's contamination. The right censoring of contaminated subjects has become standard practice in randomized clinical trials where a patient in a randomized treatment group undergoes some characteristic change that is beyond the researcher's control. Furthermore, by including right-censored observations in our data rather than truncating the data, we avoid problems caused by data truncation that could lead to spurious evidence of contagion (Van den Bulte and Iyengar 2011).[6]

We note that the exclusion of peers with multiple treated friends does not preclude measurement of network externalities. Peers of treated users that become adopters but are not connected are considered uncontaminated and are included in our analysis. For two peers of a treated user that are connected and eventually become adopters, the initial peer adopter is included in our analysis, and only the peer that subsequently adopts is considered contaminated and excluded for all times subsequent to contamination. These procedures enable a tightly controlled randomized trial of peer influence that addresses the potential for selection and leakage effects.

Robustness Checks for Different Hazard Model Specifications

There are a limited number of survival models that apply to contexts with multiple failures. Among these models, the variance-corrected stratified proportional hazards model reported in this paper is the most appropriate specification given the structure of our data and the parameters we estimate. However, we also checked multiple other hazard model specifications to test the robustness of our results to changes in model specification and estimation strategy.

[6] An alternative to dealing with users who have multiple adopter friends belonging to different treatment groups would be to ascribe peer adoption outcomes to influence from the adopter friend in the viral treatment group with the fewest viral features (above the baseline). However, we do not know a priori that inclusion of a viral feature does not have a negative impact on peer adoption outcomes (for example, if peers view viral messages as spam). To ascribe users with multiple adopter friends as peers of the lowest viral treatment group could upwardly or downwardly bias estimates of lower viral treatment effectiveness. It is therefore cleaner to remove contaminated users altogether.

Table A.1 Robustness Checks for Different Model Specifications

Specification	1 VCSPHM	2 AFT	3 EXP	4 AG_1	5 AG_2	6 PWP	7 WLW
Viral state = passive	2.51**	−2.41**	1.01***	2.60***	2.54***	2.51***	2.00*
	(0.86)	(1.16)	(0.35)	(0.91)	(0.87)	(0.865)	(0.78)
Viral state = active	3.31***	−3.66***	1.30***	3.51***	3.30***	3.31***	2.62**
	(1.24)	(1.22)	(0.39)	(1.36)	(1.26)	(1.24)	(1.02)
Application activity	1.02***	−0.119***	0.015***	1.02***	1.02***	1.02***	1.00
	(0.003)	(0.039)	(0.003)	(0.003)	(0.003)	(0.003)	(0.002)
Notifications	1.02***	−0.115***	0.025***	1.02***	1.02***	1.02***	1.01***
	(0.002)	(0.010)	(0.002)	(0.002)	(0.001)	(0.002)	(0.002)
Invites	1.06**	−0.198	0.090**	1.07*	1.06**	1.06**	1.02
	(0.028)	(0.259)	(0.036)	(0.037)	(0.035)	(0.027)	(0.018)
Prior adopters					1.50***		
					(0.062)		
Time dummies	No	Yes	Yes	No	No	No	No
Log likelihood	−4,542.58	−2,826.32	−4,136.53	−5,254.17	−5,212.88	−4,542.56	−4,561.56
χ^2 (df)	307.47***	—	1,656.60***	412.65***	435.88***	307.60***	109.17***
	(5)		(11)	(5)	(6)	(5)	(5)
Observations	3,929	3,929	3,929	3,929	3,929	3,929	3,929

Notes. Standard errors are clusters around users' local network neighborhoods. VCSPHM, variance-corrected stratified proportional hazards model as specified and reported in the paper; AFT, accelerated failure time model with log-logistic survival distribution; EXP, exponential regression with log relative-hazard form; AG, Andersen and Gill (1982) model; PWP, Prentice et al. (1981) proportional hazards model with time-dependent strata; WLW, Wei et al. (1989) marginal risk set model.

*$p < 0.10$; **$p < 0.05$; ***$p < 0.001$.

For good reviews of appropriate specifications of survival models in multiple failure data, we recommend Wei and Glidden (1997) and Ezell et al. (2003).

Table A.1 reports results of different hazard model specifications, all of which are similar to our own. We report the original variance-corrected stratified proportional hazards model specification detailed in this paper in column 1. Column 2 reports an accelerated failure time model with a log-logistic survival distribution. Column 3 reports an exponential regression with log relative-hazard form. Column 4 reports results from a traditional Andersen and Gill (1982) model. Column 5 introduces a time-dependent covariate measuring the number of prior adopters to the traditional Andersen–Gill specification to capture the dependence structure among recurrence times, which in our original model is captured by the adoption event strata k. Column 6 reports a Prentice et al. (1981) proportional hazards specification with time dependent strata. Finally, column 7 reports results of a Wei et al. (1989) marginal risk set model. We note that all specifications produce similar results. However, we are most confident in our original specification, which is best suited to our context and data.

References

Andersen, P. K., R. D. Gill. 1982. Cox's regression model for counting processes: A large sample study. *Ann. Statist.* **10**(4) 1100–1120.

Aral, S. 2011. Identifying social influence: A comment on opinion leadership and social contagion in new product diffusion. *Marketing Sci.* **30**(2) 217–223.

Aral, S., M. Van Alstyne. 2011. The diversity-bandwidth tradeoff. *Amer. J. Sociol.* Forthcoming.

Aral, S., E. Brynjolfsson, M. Van Alstyne. 2007. Productivity effects of information diffusion in e-mail networks. *ICIS 2007 Proc.*, Paper 17, http://aisel.aisnet.org/icis2007/17.

Aral, S., L. Muchnik, A. Sundararajan. 2009. Distinguishing influence-based contagion from homophily-driven diffusion in dynamic networks. *Proc. Natl. Acad. Sci. USA* **106**(51) 21544–21549.

Aral, S., L. Muchnik, A. Sundararajan. 2011. Engineering social contagions: Optimal network seeding and incentive strategies. Working paper, Stern School of Business, New York University, New York.

Berger, J., C. Heath. 2005. Idea habitats: How the prevalence of environmental cues influences the success of ideas. *Cognitive Sci.* **29**(2) 195–221.

Biyalogorsky, E., E. Gerstner, B. Libai. 2001. Customer referral management: Optimal reward programs. *Marketing Sci.* **20**(1) 82–95.

Bramoulle, Y., H. Djebbari, B. Fortin. 2009. Identification of peer effects through social networks. *J. Econom.* **150**(1) 41–55.

Brown, J. J., P. H. Reingen. 1987. Social ties and word-of-mouth referral behavior. *J. Consumer Res.* **14**(3) 350–362.

Christakis, N. A., J. H. Fowler. 2007. The spread of obesity in a large social network over 32 years. *New England J. Medicine* **357**(4) 370–379.

Coleman, J. S. 1988. Social capital in the creation of human capital. *Amer. J. Sociol.* **94**(Supplement) S95–S120.

Dellarocas, C. 2006. Strategic manipulation of Internet opinion forums: Implications for consumers and firms. *Management Sci.* **52**(10) 1577–1593.

DoubleClick. 2009. DoubleClick benchmarks report: 2009 year-in-review. White paper, DoubleClick, New York. http://www.google.com/doubleclick/research/index.html.

Duflo, E., R. Glennerster, M. Kremer. 2008. Using randomization in development economics research: A toolkit. T. P. Schultz, J. Strauss, eds. *Handbook of Development Economics*, Vol. 4, Chap. 61. North-Holland, Amsterdam, 3895–3962.

Emerson, R. M. 1962. Power-dependence relations. *Amer. Sociol. Rev.* **27**(1) 31–41.

Epsilon. 2009. Epsilon Q1 2009 email trends and benchmarks. White paper, Epsilon, Irving, TX. http://www.epsilon.com/pdf/EmailTrendandBenchmarkReport_Q1_09%20FINAL.pdf.

Ezell, M. E., K. C. Land, L. E. Cohen. 2003. Modeling multiple failure time data: A survey of variance—Corrected proportional hazards models with empirical applications to arrest data. *Sociol. Methodology* **33**(1) 111–167.

Frenzen, J., K. Nakamoto. 1993. Structure, cooperation, and the flow of market information. *J. Consumer Res.* **20**(3) 360–375.

Ghose, A., S.-P. Han. 2010. An empirical analysis of user content generation and usage behavior in mobile media. Working paper, Stern School of Business, New York University, New York.

Ghose, A., S. Yang. 2009. An empirical analysis of search engine advertising: Sponsored search in electronic markets. *Management Sci.* **55**(10) 1605–1622.

Godes, D., D. Mayzlin. 2004. Using online conversations to study word-of-mouth communication. *Marketing Sci.* **23**(4) 545–560.

Godes, D., D. Mayzlin. 2009. Firm-created word-of-mouth communication: Evidence from a field test. *Marketing Sci.* **28**(4) 721–739.

Goldenberg, J., S. Han, D. R. Lehmann, J. W. Hong. 2009. The role of hubs in the adoption process. *J. Marketing* **73**(2) 1–13.

Gross, R., A. Acquisti, J. Heinz. 2005. Information revelation and privacy in online social networks. *Proc. 2005 ACM Workshop on Privacy in the Electronic Soc.*, Association for Computing Machinery, New York.

Hartmann, W. R., P. Manchanda, H. Nair, M. Bothner, P. Dodds, D. Godes, K. Hosanagar, C. Tucker. 2008. Modeling social interactions: Identification, empirical methods and policy implications. *Marketing Letters* **19**(3) 287–304.

Heath, C., C. Bell, E. Sternberg. 2001. Emotional selection in memes: The case of urban legends. *J. Personality Soc. Psych.* **81**(6) 1028–1041.

Hill, S., F. Provost, C. Volinsky. 2006. Network-based marketing: Identifying likely adopters via consumer networks. *Statist. Sci.* **21**(2) 256–276.

Iyengar, R., C. Van den Bulte, T. W. Valente. 2011. Opinion leadership and social contagion in new product diffusion. *Marketing Sci.* **30**(2) 195–212.

Jackson, M. O. 2008. Average distance, diameter, and clustering in social networks with homophily. *Internet and Network Economics.* Lecture Notes in Computer Science, Vol. 5385. Springer, Berlin, 4–11.

Katz, E., P. F. Lazersfled. 1955. *Personal Influence.* Free Press, New York.

MailerMailer. 2010. Email marketing metrics report 2010 edition. White paper, MailerMailer, Rockville, MD. http://www.mailermailer.com/resources/metrics/2010/index.rwp.

Manchanda, P., Y. Xie, N. Youn. 2008. The role of targeted communication and contagion in product adoption. *Marketing Sci.* **27**(6) 961–976.

Manski, C. F. 1993. Identification problems in the social sciences. *Sociol. Methodology* **23** 1–56.

Mayzlin, D. 2006. Promotional chat on the Internet. *Marketing Sci.* **25**(2) 155–976.

Nam, S., P. Manchanda, P. K. Chintagunta. 2010. The effect of signal quality and contiguous word of mouth on customer acquisition for a video-on-demand service. *Marketing Sci.* **29**(4) 1–11.

Newman, M. E. J. 2003. Ego-centered networks and the ripple effect. *Soc. Networks* **25**(1) 83–95.

Oestreicher-Singer, G., A. Sundararajan. 2008. The visible hand of social networks in electronic markets. Working paper, Stern School of Business, New York University, New York.

Phelps, J. E., L. Lewis, L. Mobilo, D. Perry, N. Raman. 2004. Viral marketing or electronic word-of-mouth advertising: Examining consumer responses and motivations to pass along email. *J. Advertising Res.* **44**(4) 333–348.

Prentice, R. L., B. J. Williams, A. V. Peterson. 1981. On the regression analysis of multivariate failure time data. *Biometrika* **68**(2) 373–379.

Riley, E. 2009. Go big or go home advertising. Technical report, Forrester Research, Cambridge, MA. http://www.forrester.com/rb/Research/go_big_or_go_home_advertising/q/id/54722/t/2.

Sacerdote, B. 2001. Peer effects with random assignment: Results for dartmouth roommates. *Quart. J. Econom.* **116**(2) 681–704.

Snijders, T., C. Steglich, M. Schweinberger. 2006. Modeling the coevolution of networks and behavior. K. van Montfort, J. Oud, A. Satorra, eds. *Longitudinal Models in the Behavioral and Related Sciences*. Routledge Academic, London, 41–72.

Stephen, A. T., J. A. Berger. 2009. Creating contagious: How social networks and item characteristics combine to spur ongoing consumption and reinforce social epidemics. Working paper, Wharton School, University of Pennsylvania, Philadelphia.

Taibleson, M. H. 1974. Distinguishing between contagion, heterogeneity, and randomness in stochastic models. *Amer. Soc. Rev.* **39**(9) 877–880.

Tam, K. Y., S. Y. Ho. 2005. Web personalization as a persuasion strategy: An elaboration likelihood model perspective. *Inform. Systems Res.* **16**(3) 271–291.

Trusov, M., R. E. Bucklin, K. Pauwels. 2009. Effects of word-of-mouth versus traditional marketing: Findings from an Internet social networking site. *J. Marketing* **73**(5) 90–102.

Tucker, C. 2008. Identifying formal and informal influence in technology adoption with network externalities. *Management Sci.* **54**(12) 2024–2038.

Tucker, C. 2010. Social networks, personalized advertising, and privacy controls. Working Paper 10-07, Net Institute, New York.

Van den Bulte, C., R. Iyengar. 2011. Tricked by truncation: Spurious duration dependence and social contagion in hazard models. *Marketing Sci.* **30**(2) 233–248.

Van den Bulte, C., G. L. Lilien. 2001. Medical innovation revisited: Social contagion versus marketing effort. *Amer. J. Sociol.* **106**(5) 1409–1435.

Van den Bulte, C., S. Stremersch. 2004. Social contagion and income heterogeneity in new product diffusion: A meta-analytic test. *Marketing Sci.* **23**(4) 530–544.

Van der Lans, R., G. van Bruggen, J. Eliashberg, B. Wierenga. 2010. A viral branching model for predicting the spread of electronic word of mouth. *Marketing Sci.* **29**(2) 348–365.

Watts, D. J., P. S. Dodds. 2007. Influentials, networks, and public opinion formation. *J. Consumer Res.* **34**(4) 441–458.

WebMarketCentral. 2007. Email campaign, newsletter and banner ad click-through rates (CTR). WebMarketCentral (blog), August 14, http://webmarketcentral.blogspot.com/2007/08/email-campaign-newsletter-and-banner-ad.html.

Wei, L. J., D. V. Glidden. 1997. An overview of statistical methods for multiple failure time data in clinical trials. *Statist. Medicine* **16**(8) 833–839.

Wei, L. J., D. Y. Lin, L. Weissfeld. 1989. Regression analysis of multivariate incomplete failure time data by modeling marginal distributions. *J. Amer. Statist. Assoc.* **84**(408) 1065–1073.

Wu, F., B. Huberman, L. Adamic, J. Tyler. 2004. Information flow in social groups. *Physica A: Statist. Theoret. Phys.* **337**(1–2) 327–335.