Toward Effective Social Contagion: A Micro Level Analysis of the Impact of Dyadic Network Relationship

Completed Research Paper

Jing Peng*
University of Pennsylvania
3730 Walnut ST STE 500, PA 19104
jingpeng@wharton.upenn.edu

Ashish Agarwal
University of Texas at Austin
CBA 5.234, 2110 Speedway, TX 78712
ashish.agarwal@mccombs.utexas.edu

Kartik Hosanagar
University of Pennsylvania
3730 Walnut ST STE 500, PA 19104
kartikh@wharton.upenn.edu

Raghuram Iyengar
University of Pennsylvania
3730 Walnut ST STE 700, PA 19104
riyengar@wharton.upenn.edu

Abstract

Social advertising holds the potential to reshape the traditional advertising industry. Understanding what leads to effective social contagion at the dyadic level lies at the core of cost-effective social advertising strategies. This paper is the first attempt to comprehensively study the effect of dyadic network relationship on social contagion in directed networks. This paper reveals several intriguing findings of great importance to social media marketers: (1) reciprocal followers of adopters are less likely to be influenced than non-reciprocal followers, as moderated by the popularity and novelty of information; (2) social media users pay attention to their followers' tastes while making the adoption decision; (3) the number of common mutual followers has opposite effects on the dyadic influence between reciprocal (positive) and non-reciprocal (negative) ties. In addition, this paper provides a novel model to identify social influence when adoption events are caused by multiple sources.

Keywords: Social Contagion, Reciprocity, Embeddedness, Collective Cause

Introduction

Social advertising, which uses the social connections to reach and influence a target audience, continues to show an incredible amount of growth in the industry. The social advertising market is estimated to reach $11 billion by 2017, up from 4.7 billion in 2012 (BIA/Kelsey 2013). Three quarters of advertisers indicated they had used social media for advertising, and 64% of them planned to increase the budget on social advertising in the future (Nielsen 2013). Among all forms of ads displayed on social media platforms, native ads are becoming increasingly popular in the past few years. Native ads are ads that are seamlessly integrated into users' feeds and appear nearly indistinguishable from organic content (BI Intelligence 2013), such as sponsored stories on Facebook, promoted tweets on Twitter, and Diggable ads on Digg. It is estimated that native ads will account for at least 40% or more of the $11 billion social advertising spend by 2017 (BI Intelligence 2013). One major advantage of native ads is that their less intrusive look and feel makes users more likely to click on them. It had been reported that the Click-Through-Rate on Diggable ads was 10 times higher than non-native ads (Digg and SocialMediaGroup

2010). As compared to other socially-enabled ads, the success of native ads rely more heavily on effective social contagion, as social feeds are often the primary or even sole source of realized impressions for native ads.

Social contagion is a dyadic process in which adoption behavior or information diffuses from an adopter (sender) to a non-adopter (receiver) through some sort of network connections. A fundamental question in understanding what leads to effective social contagion is how the dyadic network relationship between the sender and the receiver impacts the influence of the sender on the receiver (Aral and Walker 2014). Two primary attributes used to characterize the dyadic network relationship between two users are reciprocity and embeddedness. Reciprocity represents whether the sender and the receiver are mutual followers of each other, which is meaningful only in directed networks. Embeddedness represents the overlap between the sender and the receiver's network members (Easley and Kleinberg 2010). A few studies on how reciprocity and embeddedness impact the dyadic influence between two connected users are emerging in the literature. Shi et al. (2014) analyzed the effect of reciprocity on the sharing of tweets and found that tweets were more likely to be shared by a non-reciprocal follower than a reciprocal follower. This finding is in sharp contrast with the theories on reciprocal behavior (Falk and Fischbacher 2006; Gouldner 1960) and strength of strong ties (Friedkin 1980; Friedkin 1982; Weenig and Midden 1991; Weimann 1983). Shi et al. (2014) speculated that the counterintuitive negative effect of reciprocity was moderated by the novelty of information to the receiver's followers, but they don't rule out other possible explanations. Therefore, whether reciprocity has a negative effect on social contagion and why it has negative effect is still an open question. In undirected networks, Aral and Walker (2014) and Bapna et al. (2012) studied how embeddedness, defined as the number of common friends between the sender and the receiver on Facebook, affect dyadic influence and trust between two users, respectively. As will be seen soon, embeddedness encapsulates more structural information in directed networks than in undirected networks, while no study has systematically looked at this potentially more fruitful problem yet. This paper aims at filling in the above mentioned gaps in directed networks and adding to the literature on the impact of reciprocity and embeddedness on social contagion, particularly in the context of ads diffusion.

Investigating the effects of dyadic network attributes on social contagion requires analysis of social diffusion data on the dyadic level. One major challenge that limits the analysis of social diffusion data on the dyadic level is that a receiver often received multiple feeds regarding the same product from different senders (or a combined feed with multiple senders) before the adoption and the contribution of each sender is often unclear. Lacking of information on the quantitative contribution of each co-sender makes it impossible to identify the effects of the sender and the receiver's characteristics in social contagion. In response to this problem, this paper proposes a novel proportional hazards model that can automatically identify the contribution of each co-sender based on her characteristics. The dyadic level study of social contagion also imposes some requirements on the richness of the data used for analysis, such as the availability of social graph information and the detailed diffusion information of products. We collected a dataset meeting such requirements from Digg.com, in which we focus on the diffusion of native ads posted on Digg. As compared to previous research findings based on the diffusion of content, our findings based on diffusion of native ads are more valuable to social media marketers.

In our analysis, we rediscovered the counterintuitive finding that reciprocity has a negative effect on adoption (Shi et al. 2014). However, our analysis suggests that it is primarily moderated by the popularity (or perceived quality) of information, rather than the novelty of information. Moreover, while previous studies all find that embeddedness has positive effect on dyadic interactions in undirected networks, the effect of embeddedness are more complicated and intriguing in directed networks. Specifically, we find that the number of common followers has positive effect on social contagion, while the number of common followees has no effect. This finding demonstrates that a user cares about her followers interest/taste while making an adoption decision. In addition, we find that the effect of common mutual followers, as a counterpart to the number of common friends in undirected networks, has positive effect on social contagion only for reciprocal followers. For non-reciprocal followers, the effect of the number of common mutual followers is negative, which can be explained by the echo hypothesis (Burt 2001; Burt 2009) largely ignored in the literature of social influence. All these findings provide novel insights into the effect of dyadic network relationship on social contagion and are of great importance to researchers and practitioners in social advertising. For easier exposition, we summarize the glossaries used in this paper in Table 1.

Glossary	Description
Followee	A user followed by the focal user
Follower	A user following the focal user
Mutual/Reciprocal follower	A user following and followed by the focal user
Reciprocity	Whether two users follow each other mutually
Embeddedness	The overlap in network members between two users
Feed	A piece of information notifying the adoption activity of a followee
Co-senders	The set of adopted followees of a receiver

Table 1. Glossaries

Theory

Reciprocity

The effect of reciprocity on social contagion rests on two streams of theories: (1) the reciprocal behavior between bidirectionally connected agents in social psychology and economics (Falk and Fischbacher 2006; Gouldner 1960) and (2) the strength of strong ties in information diffusion in sociology (Friedkin 1980; Friedkin 1982; Weenig and Midden 1991; Weimann 1983). Reciprocal behavior refers to the behavior of responding to a positive (negative) action with another positive (negative) action (Gouldner 1960). In online social networks, reciprocal behavior often shows up as retweeting/commenting/liking one another's post to ensure ongoing mutual support. Anticipated reciprocal support is an important driver of online engagements. Therefore, reciprocity is expected to have a positive effect on social contagion. Meanwhile, reciprocal ties are often stronger than non-reciprocal ties (Friedkin 1980; Granovetter 1973). In fact, many studies determine whether a tie is strong or weak simply based on whether or not it is reciprocal (Friedkin 1982; Shi et al. 2014). Since the publication of the famous strength of weak tie paper (Granovetter 1973), a large body of literature had been devoted to compare the strength of strong and weak ties in various scenarios, at both the dyadic and aggregate level. So far, a fairly robust finding across different studies is that strong ties are more effective in influencing adoption decisions due to its efficiency on the dyadic level (Weenig and Midden 1991; Weimann 1983). Whereas, weak ties are more effective in spreading information due to its structural advantage on the aggregate level (Granovetter 1973; Weimann 1983). Therefore, the literature on strength of strong ties also predicts that reciprocity has a positive effect on social contagion.

Surprisingly, according to a recent study on the sharing of Twitter content, Shi et al. (2014) find that a tweet is more likely to be shared by a non-reciprocal follower than a reciprocal follower. This finding is in sharp contrast to the theories on reciprocal behavior and strength of strong ties. Shi et al. speculate that the negative effect of reciprocity on sharing is moderated by the novelty of information to the receiver's followers. However, they don't have a metric that can directly measure the novelty of information and their results don't rule out other possible explanations. Therefore, whether reciprocity has a negative effect on social contagion and why it may have a negative effect is still an open question. The first goal of this paper is to (1) confirm whether reciprocity has a negative main effect on social contagion and (2) find out the potential variables that moderate the effect of reciprocity if (1) is true. Two potential moderating variables we are interested in are the popularity and novelty of information. The popularity of information is a signal of quality and can affect a receiver's trust on it. Given the fundamental role of trust in social contagion, the effect of popularity on social contagion is not a trivial issue. The novelty of information is also an important problem because users often see duplicated ads on the same product. Understanding users' attitude on the novelty of information can help marketers optimize the frequency of marketing campaigns on social networks.

Embeddedness

Easley and Kleinberg (2010) defined the embeddedness of a network tie between two users as the number of common neighbors they have. In undirected or bidirectional network like Facebook, there is no ambiguity that embeddedness should be defined as the number of common friends between two users. However, in directed network like Twitter, embeddedness can be defined in different ways depending on the interpretation of "neighbor". By interpreting a neighbor as a follower/followee/mutual follower, we can define embeddedness as the number of common followers/followees/mutual followers between two connected users. Each of these three embeddedness metrics provides a different angle to look at the network relationship between two users. The effects of all these three embeddedness metrics on social contagion will be studied in this paper.

Each of these metrics can have a positive impact on the dyadic diffusion process. In directed networks, people follow or subscribe to selective others to keep themselves informed about activities they are potentially interested in. The composition of one's followees largely reflects her topical interest or taste. The more common followees two users have, the more likely they have similar interest, and the more likely they will respond to feeds from each other. Likewise, the composition of one's followers reflects the topical interest of her audience. The interest of audience has great impact on the adoption decision of a user because what adopted by her will be broadcasted to her audience. According to the outcome expectation theory (Bandura 1997), a user is more likely to adopt a "product" her followers will like, as endorsing for a "product" perceived to be unsound or irrelevant can hurt her reputation (Bock et al. 2005). In fact, it has been shown that the preference of audience impacts whether one would retweet an tweet on Twitter (Leadtail 2013). The adoption of the sender is a signal that the sender believes the "product" fits the interest of her followers. Therefore, the more followers a receiver shares with the sender, the more likely the receiver will make the same decision as the sender. To sum up, a receiver cares about her own interest will be positively affected by the number of common followees with the sender, while a receivers cares about her followers' interest will be positively affected by the number of common followers with the sender.

The number of common mutual followers characterizes mutual accessibility of one to another through third-parties, which might be the most appropriate counterpart to the embeddedness defined in undirected network. Based on the type of a network tie (positive vs. negative, or friend vs. enemy), Burt proposed two hypotheses regarding how the number of common neighbors may affect dyadic trust in undirected networks (Burt 2001; Burt 2009). The first hypothesis is the bandwidth hypothesis: when a network tie between two users is positive, the existence of common neighbors expands the bandwidth of communication and makes their evaluation of each other more accurate and confident. Therefore, the level of trust between two users increases with the number of third-party ties bridging them together. The second hypothesis is the echo hypothesis: when a network tie between two users is negative, their common neighbors are more likely to disclose negative information about one to another under the pressure of social etiquette, and therefore the existence of common neighbors simply reinforces the predisposition of one on another. Clearly, these two hypotheses predict opposite directions on the effect of common neighbors on dyadic trust. If we treat a reciprocal tie as a positive tie and a non-reciprocal tie as a negative tie, the bandwidth and echo hypotheses can be easily extended to directed networks. The negative aspect of a non-reciprocal tie is that, the more common mutual followers two users have, the more problematic it is for the tie between them to be non-reciprocal, as positive ties encourage triadic closure (Friedkin 1980; Granovetter 1973; Weimann 1983). If this extension is appropriate, then the number of common mutual followers should have a positive/negative effect on the adoption of a reciprocal/non-reciprocal follower. However, the validity of the bandwidth and echo hypotheses in directed networks remain to be tested.

The study on the effect of embeddedness on social contagion is just emerging in the literature of social influence and primarily limited to directed networks. Aral and Walker (2014) found that the number of common friends on Facebook had strong positive effect on dyadic level influence. Bapna et al. (2012) found that higher embeddedness led to higher dyadic level trust while playing the trust game among Facebook users. Both findings lend support to the bandwidth hypothesis in undirected networks. In directed networks, however, no work has been done to comprehensively study the effects of the earlier mentioned three embeddedness metrics on dyadic level influence. The only work we are aware of which studies the effect of one of the three embeddedness metrics in social contagion is done by Shi et al. (2014). They found that the number of common followers had a negative effect on adoption. However, they didn't control for the numbers of common followees and common mutual followers which are highly correlated with the number of common followers, so the finding is not very convincing. The second goal of this paper is to investigate the effects of the three embeddedness metrics in directed networks in social contagion process. Given that previous studies predominantly focus on the bandwidth hypothesis, we are particularly interested in whether the echo hypothesis is valid in directed networks.

Model

In observational and experimental studies of social contagion, one common problem is that a receiver received multiple feeds on the same "product" from different senders before adopting the "product" and the contribution of each co-sender on the adoption is unclear. This problem is also known as the multi-channel attribution problem in the advertising industry (Abhishek et al. 2012). Lacking of information on the quantitative contribution of each co-sender on the adoption events make it impossible to identify social influence using the standard proportional hazards model (Cox 1972), which has become the most popular approach in the literature. The statistical challenge on attribution greatly limits the way social contagion data are analyzed in the literature, especially when dyadic level analysis is desired. For example, in a study which "analyze data from a social networking site to identify WOM effects at the individual level", the authors averaged the characteristics on the sender side to get rid of the attribution problem (Katona et al. 2011). In another study which analyzes social diffusion data on the dyadic level, the adoption events that can be potentially caused by multiple senders are deleted (Aral and Walker 2012). These methods either compromise in model precision or data precision. In response to this statistic challenge, this paper proposes a novel proportional hazards model to estimate the contribution of individual senders when multiple co-senders collectively cause an adoption.

Basic Formulation of Dyadic Hazard

We present our notations in the context of ads diffusion over social network. Let i, j, and k index senders, receivers, and ads, respectively. Let t be the time elapsed since the creation of an ad. Let $X_i(t)$ and $X_j(t)$ represent the unitary attributes of sender i and receiver j, respectively. Let X_{ij} represent the dyadic attributes concerning sender i and receiver j. Let $\lambda_{ijk}(t)$ represents the dyadic level hazard of sender i causing receiver j to adopt ad k at time t. Let $\lambda_{k0}(t)$ represents the baseline hazard for ad k. The dyadic level hazard, stratified on ads, is given by

$$\lambda_{ijk}(t) = \lambda_{k0}(t)\, exp\big(\beta_1 X_i(t) + \beta_2 X_j(t) + \beta_3 X_{ij}(t) + \varepsilon_i\big) \tag{1}$$

where $\varepsilon_i \sim N(0, \sigma)$ is the shared frailty term used to capture unobserved confounders on the sender side. Imposing a similar frailty term on the receiver side is also plausible, but unnecessary as the reoccurring frequencies of receivers are often considerably lower than the reoccurring frequencies of senders. As will be seen soon in Equation (2), the baseline hazard $\lambda_{k0}(t)$ will be cancelled out in the likelihood. Therefore, the above semi-parametric formulation allows $\lambda_{k0}(t)$ to change arbitrary over time and across ads, which frees us from the hassle of dealing with the effects of ad attributes. The above formulation of dyadic hazard is similar to the formulations given in (Aral and Walker 2012) and (Lu et al. 2013), but we allow one receiver to see the same feed from multiple senders.

Model Estimation

Let $\theta = \{\beta_1, \beta_2, \beta_3, \varepsilon_1, \dots, \varepsilon_m, \sigma\}$ represent the full set of parameters in our model. Let $R_k(t)$ represent the set of receivers who haven't adopted ad k before time t (excluding). Let $I_{jk}(t)$ represent the set of co-senders that have sent a feed regarding ad k to receiver j before time t. Let E represent the whole set of adoption events observed in the data and let E_{jk} represents the event of receiver j adopting ad k. The key assumption of the proposed proportional hazard model is that the adoption of a receiver is collectively caused by all her co-senders. In a hazards model, this means that the adoption time of a receiver is determined by the overall hazard of the receiver. This is a quite reasonable assumption as every co-sender should more or less have some effect on the receiver. Suppose event E_{jk} occurred at time τ_{jk}, under the collective cause assumption, the partial log likelihood of this event can be written as

$$l\big(E_{jk}|\theta\big) = ln\, p\big(E_{jk}|\theta\big) = ln\left(\frac{\sum_{i\in I_{jk}(\tau_{jk})} \lambda_{ijk}(\tau_{jk})}{\sum_{j'\in R_k(\tau_{jk})} \sum_{i'\in I_{j'k}(\tau_{jk})} \lambda_{i'j'k}(\tau_{jk})}\right) \tag{2}$$

where $\sum_{i\in I_{jk}(\tau_{jk})} \lambda_{ijk}(\tau_{jk})$ represents the joint hazard of receiver j to be influenced by all co-senders together. The addictive form of the joint hazard results from the conditional independence of hazards, which is a standard assumption in proportional hazards model. The overall partial likelihood of the entire dataset can be written as

$$l(E|\theta) = \sum_{E_{jk} \in E} l(E_{jk}|\theta) = \sum_{E_{jk} \in E} ln\left(\frac{\sum_{i \in I_{jk}(\tau_{jk})} \lambda_{ijk}(\tau_{jk})}{\sum_{j' \in R_k(\tau_{jk})} \sum_{i' \in I_{j'k}(\tau_{jk})} \lambda_{i'j'k}(\tau_{jk})}\right) \quad (3)$$

The parameters in our model can be estimated by maximizing the partial log likelihood given in Equation (3) using the Newton-Raphson method or other numerical optimization methods. One thing worth noting is that, in contrast to the partial likelihood of the standard proportional hazards model, the above partial likelihood is not guaranteed to be globally concave. When the information matrix is not positive definite in a search region, we modify the information matrix slightly using the approach employed by Schnabel and Eskow (1999) to ensure that the Newton-Raphson method can still work. The effectiveness of the enhanced New-Raphson algorithm has been illustrated through extensive simulations in the Appendix.

One advantage of the proposed model is that it doesn't speculate on the contribution of each co-sender apriori, but let the model automatically determine the contribution of individual co-senders based on their characteristics. In Appendix A, we have shown from simulation that when the adoptions are truly collectively-caused, the proposed model can recover the true parameters with negligible errors. Furthermore, our simulation shows the collective cause model can still give unbiased estimates if even only one of the co-senders impacts the decision of the receiver.

Modeling Spontaneous Adoptions

The basic formulation of the dyadic hazard ignores the possibility of users to adopt spontaneously. For example, one can purchase a product after seeing friends sharing their experience with this product on Facebook, or after seeing the product in a Web store. The former type of purchase is an influenced adoption, while the latter type of purchase is a spontaneous adoption. Generally speaking, any kind of adoptions caused by non-social sources can be treated as spontaneous adoptions. Teasing out spontaneous adoptions from influenced adoptions is the primary challenge for the study of social contagion. In the basic dyadic hazard formulation, each social source (friend or followee) can be represented by a specific sender. In order to incorporate the impact of non-social sources, we can treat all non-social sources as a special sender and use a dummy variable to capture the effect of this special sender. Specifically, the dyadic hazard can be rewritten as

$$\lambda_{ijk}(t) = \lambda_{k0}(t)\, exp\big(\beta_0 s_i + \beta_1 X_i(t) + \beta_2 X_j(t) + \beta_3 X_{ij}(t) + \varepsilon_i\big) \quad (4)$$

where the dummy variable s_i indicates whether the sender is the special sender. For the special sender, all undefined unitary and dyadic attributes are coded as missing and set to zero (or any other default values as the selection of default values only affects β_0). β_0 captures the combined effect of all non-social sources, as compared to a normal sender with attributes valued at zero, on the adoption of the receiver. In fact, β_0 cannot be identified if we impose shared frailty terms on senders, as it will be absorbed into the frailty term of the special sender (i.e., ε_0). Given that the special sender is intrinsically different from normal senders, we allow the variance of the frailty term for the special sender to be different from normal senders. Since everyone can adopt spontaneously, the special sender is a co-sender for every potential adopter by default. This dummy variable setup enables us to seamlessly incorporate the effect of non-social sources without affecting the way the proposed model is estimated.

For the purpose of this particular study, we further complicate the dyadic hazard by introducing feed attributes $X_{ik}(t)$, such as the daypart of the feed and the adoption time of the sender (i.e., the time it takes for the sender to adopt since the creation of the ad). Excluding interaction terms, the final dyadic hazard for our study is given by

$$\lambda_{ijk}(t) = \lambda_{k0}(t)\, exp\big(\beta_0 s_i + \beta_1 X_i(t) + \beta_2 X_j(t) + \beta_3 X_{ij}(t) + \beta_4 X_{ik}(t) + \varepsilon_i\big) \quad (5)$$

Interaction terms can be easily added into the model as needed. It will be pretty straightforward to see which interactions have been considered in our later analysis, so we don't discuss them in detail here.

Data

The dyadic level study of social contagion imposes some requirements on the richness of the data used for analysis, such as the availability of social graph information and the detailed diffusion information of products. We collected a dataset meeting such requirement from Digg.com, one of the largest online social news aggregating websites. As a social news aggregator, Digg does not produce any content itself but encourages users to submit and vote for stories of their interests. The collective votes (called diggs or digg number) on a story play a critical role in determining whether the story can be featured on the front page or not. Digg maintains a Twitter alike social network structure, allowing users to follow each other. On Digg, users can explore top stories in the front page or navigate through feeds about their followees' activities (digging, commenting, and submitting stories) in the "My News" page. Digg introduced a native advertising model, called diggable ads, in 2009, which were not removed from the website until Digg was acquired by another company in August 2012. Diggable ads were seamlessly integrated into organic stories and displayed at three fixed positions of the eighteen slots available on the front page. Diggable ads were identical to organic stories except for an inconspicuous flag "sponsored by xx" below them. Users can digg up or down a diggable ad just like digging an organic story. Diggable ads are removed from the front page when running out of budget, but users can still see them from social feeds.

We collected all diggable ads appeared on the website between October 2010 and July 2012, including the basic information of those ads (e.g., title, description, URL, and time of creation) and the activities (i.e., diggs and comments) on them. We also collected the profile information of all users who ever dugg one of the diggable ads, including gender, location, number of diggs, number of comments, number of submissions, number of followers, and number of followees, as of June/7/2012. Moreover, we spent 19 days (June/7/2012- June/26/2012) to collect the complete set of followers and followees for these users. For the purpose of this study, we focus on 31 ads created between May 24 and June 25 to ensure that the real network structure does not deviate much from the snapshot we had collected. Most of the ads were posted on three days: May 24, June 1, and June 25. We selected all 7,709 users who had dugg at least one ad before and were still active on Digg in the focal time period as potential adopters. This allows us to follow on users who have real chances to adopt ads.

On Digg, the front page is the primary non-social sources for adoption of ads. For easier exposition, we refer to the special sender, used to represent all non-social sources, as the front page. For each potential adopter, we generate one dyadic observation for her if one of her followees adopts, until she adopts. Since everyone has access to the front page, we generate one additional dyadic observation for each potential adopter, with the front page being the sender. A user is considered an adopter if she diggs or comments on the ad. One adoption is considered spontaneous if none of her followees adopt the ad before her. Otherwise, it is considered as a potential influenced adoption. One converts from a receiver to a sender immediately after the adoption, implying that senders are a subset of receivers. Since the number of comments is negligible as compared to the number of diggs, we treat comments indifferently as diggs. The summary statistics of the final dataset is shown in Table 2. 66.6% of adoptions have more than one co-sender, and the average number of co-senders is 3.46, including the front page.

Number of ads	31
Number of adopters (senders)	1802
Number of potential adopters (receivers)	7709
Number of <sender, receiver, ad> tuples	560,912
Tuples in which senders are the front page	238,979 (42.6%)
Tuples in which senders are followees	321,933 (57.4%)
Number of adoptions (diggs and comments)	2231
Number of diggs	2229
Number of comments	2
Number of spontaneous adoptions	745 (33.4%)
Number of potential influenced adoptions	1486 (66.6%)
Average number of co-senders per adoption	3.46

Table 2. Description of Cleaned Dataset

In Table 3, we summarize the statistics of unitary network attributes of potential adopters and the dyadic network attributes for observed <sender, receiver> dyads. Table 3 also includes the diffusion statistics of ads.

	Mean	St. dev.	Min	Median	Max
Unitary Network Attributes of Potential Adopters					
Number of followees	309.7	454.6	1	149	10122
Number of followers	449.3	1427.4	0	161	94235
Number of mutuals	132.2	219.4	0	49	4600
Dyadic Network Attributes of All Sender-Receiver Dyads					
isMutual	1.3	0.4	1	1	2
Number of common followees	59.0	74.4	0	33	1105
Number of common followers	121.6	343.3	0	39	10153
Number of common mutuals	17.8	36.2	0	5	594
Ad Attributes					
Number of diggs on ads	102.5	81.7	14	98	295

Table 3. Key Statistics of Cleaned Dataset

To effectively estimate the effects of dyadic network attributes, we have controlled the demographics of users, the engagement levels of users, the unitary network attributes of users, and the attributes of feeds. Table 4 summarizes the independent variables used in our analysis.

Independent Variable		**Description**
X_i/X_j		**Attributes of sender i / receiver j**
Network attributes	followees	Number of followees (out-degree)
	followers	Number of followers (in-degree)
	mutuals	Number of mutual followers
Engagement levels	diggs	Total number of diggs
	comments	Total number of comments
	submissions	Total number of submissions
	avgDiggs	Average number of diggs per month
	avgComments	Average number of comments per month
	avgSubmissions	Average number of submissions per month
Demographics	gender	Male, female, or missing
	regMonths	How many months has the user been registered on the website
X_{ij}		**Attributes of an sender-recipient pair**
Dyadic network attributes	isMutual	Are sender i and recipient j mutual friends
	commonFollowees	Number of followees followed by both the sender and receiver
	commonFollowers	Number of followers following both the sender and receiver
	commonMutuals	Number of mutual followers shared by the sender and the receiver
X_{ik}		**Attributes of a feed**
Feed attributes	wday	Day of a week when sender i dugg ad k
	hour	Hour of a day when sender i dugg ad k
	adoptionTime	Hours taken for sender i to adopt since the creation of ad k, 0 for the front page
Miscellaneous		
	diggNum	Number of diggs on an ad at a given time point
	adoptedFollowees	Number of adopted followees in the receiver's network
	isSocial (s_i)	1 if sender i is a social source (i.e., followee), otherwise 0

Table 4. Descriptions of Independent Variables

Results

We compared a rich set of model specifications for our analysis. Table 5 summarizes the key results, in which the coefficients on most control variables are omitted. Complete results on the full set of variables and the full set of model specifications are provided in Appendix B.

Reciprocity

As can be seen in Model 1 in Table 5, the main effect of reciprocity is negative and significant, which echoes the earlier finding by Shi et al. (2014) but contradicts the theories on reciprocal behavior (Falk and

Fischbacher 2006; Gouldner 1960) and strength of strong ties (Friedkin 1980; Friedkin 1982; Weenig and Midden 1991; Weimann 1983). To reconcile our empirical finding with theoretical prediction, we investigated two variables that can potentially moderate the effect of reciprocity on dyadic influence (i.e., variables that have negative interaction effects with reciprocity). The first variable is the number of diggs on an ad (i.e., logDiggNum), which represents the popularity of the ad. In a collective voting system like Digg, popularity is probably the most important signal of quality. Therefore, the popularity of an ad can affect the perceived quality of the ad by users, though its intrinsic quality remains the same over time. The reason why popularity may interact with reciprocity is that reciprocal followers should be more tolerant to unpopular ads (with low perceived quality) than non-reciprocal followers due to their stronger trust on the sender. It can be observed from Model 2 that popularity indeed negatively interacts with reciprocity, demonstrating that reciprocal followers are more likely to digg low quality ads than non-reciprocal followers. Moreover, the main effect of reciprocity becomes positive after accounting for the interaction between popularity and reciprocity, which confirms the moderating role of popularity on reciprocity. Note that the main effect of logDiggNum cannot be identified as everyone sees the same digg number for a given ad at a given time point, the effect of which will be cancelled out automatically in the likelihood.

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Reciprocity						
isMutualTrue	-0.268**	2.1588***	0.4059**	2.1141***	-0.716***	1.511***
Embeddedness						
logCommonFollowees	0.041	-0.0112	0.0537	0.0086	0.037	-0.003
logCommonFollowers	0.588***	0.6223***	0.6387***	0.6537***	0.608***	0.677***
logCommonMutuals	-0.363***	-0.3404***	-0.3676***	-0.3533***	-0.448***	-0.456***
Interactions						
isMutualTrue:logDiggNum		-0.5910***		-0.4696***		-0.467***
isMutualTrue: adoptedFollowees			-0.2188***	-0.1629***		-0.185***
isMutualTrue:logCommonMutuals					0.210***	0.291***
Controls						
adoptedFollowees	-0.092***	-0.0833***	-0.0675***	-0.0685***	-0.091***	-0.066***
adoptionTime	0.080*	0.1137**	0.0800.	0.1043*	0.072.	0.097*
…						
logLikelihood	-16838	-16825	-16828	-16821	-16835	-16816
AIC	33754	33731	33736	33725	33750	33717

* Significance levels: p<0.001 (***), p<0.01 (**),p<0.05 (*), and p<0.1 (.).

Table 5. Parameters Estimates of Different Model Specifications

The second variable is the number of adopted followees (i.e., adoptedFollowees) in the receiver's network (or number of co-senders in other words), which reflects the number of (potential) times the receiver has been exposed to the ad and hence the novelty of the ad to the receiver. Had a reciprocal follower not digging an ad after seeing it many times, the more likely this user has a problem with the ad than a non-reciprocal follower user the same condition, as the reciprocal follower is expected to respond sooner if she likes the ad, due to anticipated reciprocity and stronger trust on the sender. In this sense, the marginal effect of an additional co-sender is expected to be weaker on a reciprocal follower than a non-reciprocal follower. It can be observed in Model 3 that the interaction between adoptedFollowees and reciprocity is negative, which confirms our conjecture that the marginal effect of an additional co-sender is weaker for reciprocal followers than non-reciprocal followers. Note that the negative main effect of adoptedFollowees only suggests that the dyadic susceptibility of a receiver to an individual sender decreases with the number of co-senders, but not the overall susceptibility to all co-senders. Therefore, the negative main effect of adoptedFollowees in our dyadic model does not conflict with the earlier finding that the overall susceptibility of a user increases with the number of adopted friends in her network (Bapna and Umyarov 2012; Katona et al. 2011).

Although both logDiggNum and adoptedFollowees can moderate the main effect of reciprocity from negative to positive, it can be seen that the moderating effect of logDiggNum is stronger than adoptedFollowees as the effect size of reciprocity in Model 2 is substantially larger than that in Model 3. This finding demonstrates that the negative effect of reciprocity is primarily moderated by the popularity or the perceived quality of information, rather than the novelty of information as suggested by Shi et al. (2014).

Embeddedness

As discussed earlier, the number of common followees reflects the similarity between the sender and the receiver's interests, while the number of common followers reflects the similarity between their audiences' interests. The effect of logCommonFollowers is positive and significant, demonstrating that a user cares about her followers' interest while digging an ad. This is consistent with the earlier finding that the preference of followers impacts whether one would retweet an tweet on Twitter (Leadtail 2013). In contrast, the effect of logCommonFollowees is very close to zero and not significant, which either suggests that a user doesn't care about her own taste while digging an ad or the number of common followees doesn't capture much about two users' similarity in interests. The former interpretation is less likely to be true as digging based on one's own preference is a core spirit of Digg users.

The main effect of logCommonMutuals is consistently negative and significant across all model specifications. This finding is in sharp contrast to previous work in directed networks, which have demonstrated that the number of common friends has positive effect on dyadic trust (Bapna et al. 2012) and dyadic influence (Aral and Walker 2014). However, if we look at the interaction of logCommonMutuals and reciprocity in Models 5 and 6, we can see that the number of common mutual followers indeed has a positive effect on reciprocal followers, which confirms the validity of the bandwidth hypothesis (Burt 2001; Burt 2009) for reciprocal ties. Meanwhile, the effect of logCommonMutuals is negative for non-reciprocal followers, which confirms the validity of the echo hypothesis (Burt 2001; Burt 2009) for non-reciprocal ties. The reason that the overall effect of logCommonMutuals is negative is that the majority followers of a user are non-reciprocal. The bandwidth hypothesis or likewise theory has become the default theoretical foundation while studying network embeddedness, whereas the importance of the echo hypothesis has been largely ignored in the literature. Further research on social contagion should pay more attention to the role of the echo hypothesis, especially in directed networks.

As a side note, the positive effect of adoptionTime indicates that the most recent co-sender has the strongest impact on the adoption decision of the receiver. This finding is consistent with our intuition, demonstrating the effectiveness of the proposed model to automatically estimate the quantitative contribution of co-senders based on their characteristics.

Robustness Check

The Growth of Network Structure

One concern on the validity of our analysis is that the network structure among users grows over time but we only captured a static snapshot. The direct consequence of the inaccurate network structure information is that the number of observed co-senders for a receiver could be larger or smaller than the actual number of co-senders for the receiver, depending on whether the receiver dugg the ad before or after the time her network information were collected by us. In our dataset, almost all the ads were posted on three days: May 24, June 1, and June 25. In order to test the robustness of our results to this problem, we split the dataset into two subsets: one focuses on ads created between May 24 and June 1, and another focuses on ads created on June 25. Given that users often establish new ties but rarely break old ties, the number of co-senders is likely to be overestimated on the first dataset as the network structure is collected afterwards. On the second dataset, the number of co-senders is likely to be underestimated as most of the digging activities take place after the network structure is collected. If overestimation or underestimation of co-sender number causes major bias on our estimates, the results on these two subsets should be very different from that on the full dataset. Tables 6 and 7 summarize the results on the two subsets, respectively.

It can be observed that the estimates in Tables 6 and 7 are consistent with the estimates in Table 5, except that the effect of logCommonFollowees is positive and significant in Table 6. The positive effect of logCommonFollowees is indeed expected, but unfortunately only observed in the May24-June1 subset. We believe the effect of the number of common followees is a problem that needs further examination in future work, especially in diffusion data on other networks. Overall, the results in Tables 6 and 7 demonstrate that our results are robust to the change of network structure over time. In fact, we have also tested the robustness of our model to over-count and under-count of co-senders through simulation. The relative errors barely go beyond 5% even when the under-count or over-count problem is far more severe in the simulation data than in the Digg dataset.

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Reciprocity						
isMutualTrue	-0.222.	1.9094***	0.553**	1.840***	-0.547*	1.377***
Embeddedness						
logCommonFollowees	0.208*	0.1668	0.254*	0.216*	0.197*	0.192.
logCommonFollowers	0.514***	0.5350***	0.585***	0.583***	0.530***	0.599***
logCommonMutuals	-0.394***	-0.3793***	-0.410***	-0.400***	-0.460***	-0.487***
Interactions						
isMutualTrue:logDiggNum		-0.4800***		-0.321**		-0.315**
isMutualTrue:adoptedFollowees			-0.245***	-0.204***		-0.220***
isMutualTrue:logCommonMutuals					0.153*	0.231**
Controls						
adoptedFollowees	-0.099***	-0.0917***	-0.069***	-0.070***	-0.098***	-0.068***
imprLag	0.027	0.0666	0.007	0.041	0.024	0.042
…						
logLikelihood	-11425	-11419	-11417	-11414	-11424	-11412
AIC	22928	22917	22913	22912	22928	22909

Table 6. Parameters Estimates on the May24-June1 Subset

	Model1	Model2	Model3	Model4	Model5	Model6
Reciprocity						
isMutualTrue	-0.58**	2.85***	-0.13	3.10***	-0.852*	2.411***
Embeddedness						
logCommonFollowees	-0.11	-0.15	-0.14	-0.17	-0.106	-0.162
logCommonFollowers	0.68***	0.68***	0.71***	0.71***	0.698***	0.767***
logCommonMutuals	-0.49***	-0.40***	-0.46***	-0.38***	-0.536***	-0.504***
Interactions						
isMutualTrue:logDiggNum		-0.93***		-1.02**		-1.043***
isMutualTrue:adoptedFollowees			-0.14**	-0.12*		-0.168*
isMutualTrue:logCommonMutuals					0.153	0.428***
Controls						
adoptedFollowees	-0.10***	-0.09***	-0.08***	-0.08***	-0.097***	-0.081***
adoptionTime	1.21	1.35	0.93	1.14	1.176	0.975
…						
logLikelihood	-5363	-5356	-5361	-5354	-5362	-5352
AIC	10801	10789	10799	10791	10803	10787

Table 7. Parameters Estimates on the June 25 Subset

Potential Multi-colinearity Problem

Another concern on the validity of our findings is the potential multi-colinearity problem among the dyadic network attributes characterizing reciprocity and embeddedness may affect the reliability of our estimates. The number of common followees, followers, and mutuals are often highly correlated with each other, partially due to the fact that common mutuals are the interaction of common followees and common followers. One way to alleviate this problem is to exclude common mutuals from both common followees and followers. Tables 8 shows the correlations of the dyadic network attributes before and after excluding common mutuals from common followees and followers. It can be observed that, the correlation between the three embeddedness metrics are 30~40% lower after excluding common mutuals.

If the high correlation among dyadic network attributes indeed causes multi-colinearity problem in our estimation, the estimates from our model should change substantially if we use the set of new dyadic network attributes in which common mutuals are excluded from common followees and followers. As can be seen from Table 9, while the effect sizes of our estimates change somehow, the directions are completely consistent with Table 5. Therefore, the validity of our analysis is not undermined by the potential multi-colinearity problem. Note that the impact of high correlation between independent variables on the estimation of proportional hazards model (nonlinear) is very different to its effect on estimation of linear models. The reason is that what matters most for the estimation of the proportional hazards model is the small portion of observations on which events are observed. Therefore, a level of correlation that may result in the multi-colinearity problem for linear models doesn't necessary results in the multi-colinearity problem for our model. In the presence of multi-colinearity problem, the parameter estimates tend to be highly unstable across different model specifications, which is not observed in our analysis.

		isMutual	logCommonFollowees	logCommonFollowers	logCommonMutuals
Including Mutuals	isMutual	1.00	0.30	0.23	0.52
	logCommonFollowees		1.00	0.70	0.69
	logCommonFollowers			1.00	0.67
	logCommonMutuals				1.00
Excluding Mutuals	isMutual	1.00	-0.32	-0.40	0.47
	logCommonFollowees		1.00	0.52	-0.42
	logCommonFollowers			1.00	-0.40
	logCommonMutuals				1.00

Table 8. Correlation of Dyadic Network Characteristics

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Reciprocity						
isMutualTrue	-0.327**	2.242***	0.301.	2.204***	-0.752***	1.643***
Embeddedness						
logCommonFollowees	-0.036	-0.076	-0.015	-0.057	-0.059	-0.089
logCommonFollowers	0.436***	0.460***	0.474***	0.483***	0.469***	0.535***
logCommonMutuals	-0.127**	-0.109*	-0.123**	-0.111*	-0.198***	-0.211***
Interactions						
isMutualTrue:logDiggNum		-0.636***		-0.545***		-0.579***
isMutualTrue:adoptedFollowees			-0.198***	-0.130***		-0.146***
isMutualTrue:logCommonMutuals					0.196**	0.305***
Controls						
adoptedFollowees	-0.087***	-0.079***	-0.066***	-0.067***	-0.085***	-0.062***
adoptionTime	0.06	0.090*	0.056	0.077.	0.05	0.065
…						
logLikelihood	-16862	-16849	-16854	-16846	-16860	-16842
AIC	33801	33778	33788	33777	33799	33769

Table 9. Parameters Estimates after Excluding Common Mutuals from Common Followees/Followers

Conclusions

Social advertising holds the potential to reshape the traditional advertising industry. Understanding what leads to effective social contagion at the dyadic level lies at the core of cost-effective social advertising strategies. While the effects of unitary network attributes have been well-studied in the literature, studies on the effects of dyadic network attributes in social contagion are just emerging and predominantly focus on undirected networks. This paper is the first attempt to comprehensively study the effect of dyadic network relationship on social contagion in directed networks. The micro-level analysis conducted in this paper not only sheds some insights into the dyadic level mechanism governing the social contagion process, but also provides useful guidelines for managers to improve the cost-effectiveness of their social marketing campaigns. In particular, as compared to previous studies which primarily focus on the diffusion of organic content in social networks, the analysis of this paper is based on the diffusion of native ads, which makes our findings more valuable to marketers.

This paper successfully reconciles the discrepancy between theoretical prediction and empirical finding on the effect of reciprocity in social contagion. Specifically, we find that the negative effect of reciprocity is moderated by the popularity and novelty of ads. As compared to non-reciprocal ties, reciprocal ties are more suitable for the diffusion of products which haven't established their reputation in the market yet. In addition, we find that a user pays attention to her followers' taste while making the adoption decision, which hasn't receive enough attention in the literature yet. More importantly, we demonstrate that the echo hypothesis which has largely been ignored in the literature does play an important role in the social contagion process in directed network. Since the number of common mutual followers has positive and negative effects on reciprocal and non-reciprocal followers, respectively, the overall effect of common mutual followers is indeed negative as non-reciprocal followers substantially outnumber reciprocal followers. All these findings are of important implications to social media marketers. Finally, this paper provides a novel tool to identify social influence when adoption events are caused by multiple sources.

Our work certainly has its limitations. One limitation is the lack of impression information. The two-stage information consumption model (Shi et al. 2014; Weenig and Midden 1991) suggests that one first reads a feed and then decides whether or not to adopt it. Without the impression information, we are essentially modelling the overall hazard of one to read and adopt an ad. This coarse modelling structure may increase the standard errors of our estimates. However, in many scenarios such as buzz marketing campaigns, advertisers can only measure reach and engagements but not impressions, in which case the approach employed in this paper might be desirable. Another limitation is that the adoptions could be driven by unobserved characteristics of the sender, the receiver, or the dyad. In the presence of such unobserved characteristics, the conditional independence assumption on dyadic observations may not hold and the parameter estimates of our model can be biased. The potential dependence between dyadic observations is mostly likely due to the unobserved characteristics on senders, as senders re-occur substantially more frequently than receivers and dyads in the dyadic observations. We use shared frailty terms on senders to account for the unobserved characteristics of senders. The results show that including frailty terms on senders has no major impact on our estimates and findings. However, the shared frailty model may not be enough if the unobserved characteristics are correlated with the network characteristics of interest, as it assumes that the unobserved characteristics are independent of observed characteristics.

Reference

Abhishek, V., Fader, P., and Hosanagar, K. 2012. "The Long Road to Online Conversion: A Model of Multi-Channel Attribution," Working Paper.

Aral, S., and Walker, D. 2012. "Identifying Influential and Susceptible Members of Social Networks," Science (337:6092), pp. 337-341.

Aral, S., and Walker, D. 2014. "Tie Strength, Embeddedness, and Social Influence: A Large-Scale Networked Experiment," Management Science (60:6), pp. 1352 - 1370.

Bandura, A. 1997. Self-Efficacy: The Exercise of Control. Macmillan.

Bapna, R., Gupta, A., Rice, S., and Sundararajan, A. 2012. "Trust, Reciprocity and the Strength of Social Ties: An Online Social Network Based Field Experiment," Working Paper.

Bapna, R., and Umyarov, A. 2012. "Do Your Online Friends Make You Pay? A Randomized Field Experiment in an Online Music Social Network," NBER Working Paper.

Bender, R., Augustin, T., and Blettner, M. 2005. "Generating Survival Times to Simulate Cox Proportional Hazards Models," Statistics in Medicine (24:11), pp. 1713-1723.

BI Intelligence. 2013. "The Rise of Native: Why Social Media Advertising Is Going in-Stream," Business Insider Intelligence.

BIA/Kelsey. 2013. "U.S. Social Ad Revenues to Reach $11b in 2017."

Bock, G.-W., Zmud, R.W., Kim, Y.-G., and Lee, J.-N. 2005. "Behavioral Intention Formation in Knowledge Sharing: Examining the Roles of Extrinsic Motivators, Social-Psychological Forces, and Organizational Climate," MIS Quarterly (29:1), pp. 87-111.

Burt, D.R. 2001. "Bandwidth and Echo: Trust, Information, and Gossip in Social Networks," in Networks and Markets: Contributions from Economics and Sociology.

Burt, R.S. 2009. Structural Holes: The Social Structure of Competition. Harvard university press.

Cox, D.R. 1972. "Regression Models and Life Tables," JR Stat Soc B (34:2), pp. 187-220.

Digg, and SocialMediaGroup. 2010. "Best Practices in Online Conversational Marketing," digg.com & socialmediagroup.com.

Easley, D., and Kleinberg, J. 2010. "Networks, Crowds, and Markets," Cambridge Univ Press (6:1), p. 6.1.
Falk, A., and Fischbacher, U. 2006. "A Theory of Reciprocity," Games and Economic Behavior (54:2), pp. 293-315.
Friedkin, N. 1980. "A Test of Structural Features of Granovetter's Strength of Weak Ties Theory," Social Networks (2:4), pp. 411-422.
Friedkin, N.E. 1982. "Information Flow through Strong and Weak Ties in Intraorganizational Social Networks," Social networks (3:4), pp. 273-285.
Gouldner, A.W. 1960. "The Norm of Reciprocity: A Preliminary Statement," American Sociological Review (25:2), pp. 161-178.
Granovetter, M.S. 1973. "The Strength of Weak Ties," American Journal of Sociology (78:6), pp. 1360-1380.
Katona, Z., Zubcsek, P.P., and Sarvary, M. 2011. "Network Effects and Personal Influences: The Diffusion of an Online Social Network," Journal of Marketing Research (48:3), pp. 425-443.
Leadtail. 2013. "B2b Social Marketing Report: Marketing Executives on Twitter."
Lu, Y., Jerath, K., and Singh, P.V. 2013. "The Emergence of Opinion Leaders in a Networked Online Community: A Dyadic Model with Time Dynamics and a Heuristic for Fast Estimation," Management Science (59:8), pp. 1783-1799.
Nielsen. 2013. "Paid Social Media Advertising - Industry Update and Best Practices."
Schnabel, R.B., and Eskow, E. 1999. "A Revised Modified Cholesky Factorization Algorithm," SIAM Journal on Optimization (9:4), pp. 1135-1148.
Shi, Z., Rui, H., and Whinston, A.B. 2014. "Content Sharing in a Social Broadcasting Environment: Evidence from Twitter," MIS Quarterly (38:1), pp. 123-142.
Weenig, M.W., and Midden, C.J. 1991. "Communication Network Influences on Information Diffusion and Persuasion," Journal of Personality and Social Psychology (61:5), p. 734.
Weimann, G. 1983. "The Strength of Weak Conversational Ties in the Flow of Information and Influence," Social Networks (5:3), pp. 245-267.

Appendix A: Simulation

The proposed model is called a collective cause model because it rests on the assumption that the adoption is caused by all co-senders collectively. This section tests the performance of the proposed collective cause model in recovering the true parameters when the data are generated under the collective cause assumption. In practice, it is also possible that only part of the co-senders contributes to the adoption. To demonstrate the effectiveness of the collective cause model in dealing with such data, we focus on an extreme case in which the adoption is caused by one of the co-senders independently (called single-cause data). With a goal to generate a dataset with 10K adoption events, we construct the single-cause dataset as follows:

1) Generate 200 senders, each with three covariates: one normal, one binomial, and one exponential
2) Generate 5,000 receivers, each with three covariates: one normal, one binomial, and one exponential
3) Randomly sample 10,000 senders and 10,000 receivers with replacement from the pool of 200 senders and 5,000 receivers, respectively. A one-to-one mapping between the 10K senders and 10K receivers results in 10K dyadic observations.
4) Randomly sample another 2,000 senders with replacement from the pool of 200 senders and map each of them to one of the 10K receivers in step (3) randomly. Those matched receivers in this step will therefore have multiple senders.
5) For each dyadic observation, compute the dyadic hazard, assuming the baseline hazard and all model parameters equal to 1.
6) Assume some distribution (e.g., exponential, weibull, and Gompertz) on the survival times and simulate a survival time for each of the 12K dyadic observations. See (Bender et al. 2005) for more details on how to do this. This leads to 12K survival times in the end.
7) If a receiver has multiple senders, choose the minimum survival time from all the co-senders as the observed survival time for all co-senders.
8) Choose the lower 20% quantile of all the survival times as the censoring time, and set the observations for which the survival time is greater than the censoring time as censored.

For the collective-cause dataset, the data generation procedures are exactly the same as above, except that steps (6) and (7) are replaced by:

6) For each of the 10K receivers, compute her aggregated hazards by summing up the hazards from all her co-senders. Assume some distribution (e.g., exponential, weibull, and Gompertz) on the survival times and simulate a survival time for each receiver based on her aggregated hazard. This leads to 10K survival times in the end.
7) If a receiver has multiple senders, set the survival time of all the dyadic observations regarding these co-senders to be the survival time of the receiver generated in step (7).

We use a dyadic setup to ensure that the structure of the simulated dataset is similar to the structure of dyadic social contagion dataset. In addition, we assume three entirely different distributions on the independent variables to test the robustness of the collective cause model to the distribution of independent variables. Moreover, we censored 80% of events to test the effectiveness of the single cause model on incomplete observations. We compared three models in total, as shown in Table A1. The Tied Events model treats the multiple dyadic observations on the same receiver as tied events and splits a unit case weight evenly among these observations, then uses the standard tie handling approaches in proportional hazard model to fit the data. The Equal Prob. model assumes that every co-senders of a receiver has equal probability to cause the adoption and then maximizes the expected partial likelihood. Table A1 summarizes the relative errors (i.e., $\frac{\hat{\beta} - \beta}{\beta}$) of different models on two simulated datasets, averaged over 20 runs. The prefix "r" indicates covariates on the receiver side. Enclosed in parentheses are the standard deviations of the relative errors. Results are robust to censoring, scaling, distribution of survival times, and average number of co-senders on a receiver.

	Single-Cause Data			**Collective-Cause Data**		
	Tied Events	Equal Prob.	Collective Cause	Tied Events	Equal Prob.	Collective Cause
normal	-0.1883 (0.02)	-0.2102 (0.03)	0.0064 (0.02)	-0.3312 (0.02)	-0.2122 (0.03)	0.0014 (0.02)
binomial	-0.1787 (0.04)	-0.1996 (0.05)	0.0060 (0.05)	-0.3236 (0.04)	-0.2079 (0.04)	-0.0053 (0.04)
exponential	-0.1539 (0.01)	-0.1776 (0.01)	-0.0006 (0.02)	-0.2729 (0.02)	-0.1766 (0.01)	-0.0003 (0.02)
rnormal	-0.1897 (0.02)	-0.2109 (0.02)	0.0030 (0.02)	-0.3254 (0.02)	-0.2079 (0.02)	0.0066 (0.02)
rbinomial	-0.1811 (0.05)	-0.2045 (0.05)	-0.0034 (0.05)	-0.3200 (0.05)	-0.2088 (0.05)	-0.0064 (0.04)
rexponential	-0.1541 (0.01)	-0.1758 (0.01)	0.0015 (0.01)	-0.2740 (0.01)	-0.1743 (0.01)	0.0022 (0.02)

Table A1. Relative Errors of the Collective Cause Model

As can be seen, the proposed collective cause model can recover the true parameters with negligible errors not only on the collective-cause data, but also on the single-cause data. This finding demonstrates that the collective cause model is still a valid model even if only part of the co-senders contributes to the adoption. The mathematical proof regarding why the collective cause model can still recover the true parameters when only one of co-senders contributes to the adoption is beyond the scope of this paper. The intuition behind this finding is that, in the single-cause data, the overall hazard of a receiver given in the numerator of equation (2) can be reinterpreted as the overall hazard of the receiver to be influenced by any single source she has seen. In this sense, the collective cause model is a truthful representation of the single cause data, except that it doesn't use the true cause information. The estimates of the tied events model and equal prob. model are both substantially biased downwards, which demonstrates that ad hoc modifications of standard proportional hazards model cannot address the attribution problem when there are multiple potential causes for one adoption event.

Appendix B

Table B1 displays the complete parameter estimates on a much richer set of model specifications than Table 5. The prefix "r" indicates attributes of the receiver. As can be seen, the findings presented in the main paper are particular robust to model specifications.

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8	Model 9	Model 10
isSocialTRUE	2.932***	3.0414***	3.147***	2.948***	3.162***	3.0462***	2.8658***	2.9164***	3.373***	3.156***
logFollowees	0.03	0.0044	0.123*	0.04	0.052	0.0732	0.0881	0.0886	0.064	0.11
logFollowers	-0.455***	-0.4841***	-0.702***	-0.464***	-0.856***	-0.9021***	-0.9264***	-0.9369***	-0.879***	-0.948***
logMutual	0.087	0.0612	0.003	0.068	0.218*	0.2702**	0.2231*	0.2639**	0.197*	0.232*
rLogFollowees	-0.148***	-0.1568***	-0.129***	-0.145***	-0.151***	-0.1466***	-0.1450***	-0.1435***	-0.156***	-0.150***
rLogFollowers	-0.251***	-0.2541***	-0.293***	-0.252***	-0.322***	-0.3285***	-0.3294***	-0.3326***	-0.324***	-0.336***
rLogMutual	0.086**	0.0840**	0.067*	0.081**	0.110***	0.1125***	0.1090***	0.1123***	0.114***	0.119***
isMutualTrue	-0.051	-0.1169	-0.311**	-0.073	-0.268**	2.1588***	0.4059**	2.1141***	-0.716***	1.511***
logCommonFollowees		0.1192**			0.041	-0.0112	0.0537	0.0086	0.037	-0.003
logCommonFollowers			0.335***		0.588***	0.6223***	0.6387***	0.6537***	0.608***	0.677***
logCommonMutuals				0.032	-0.363***	-0.3404***	-0.3676***	-0.3533***	-0.448***	-0.456***
isMutualTrue:logDiggNum						-0.5910***		-0.4696***		-0.467***
isMutualTrue:adoptedFollowees							-0.2188***	-0.1629***		-0.185***
isMutualTrue:logCommonMutuals									0.210***	0.291***
adoptedFollowees	-0.081***	-0.0881***	-0.090***	-0.082***	-0.092***	-0.0833***	-0.0675***	-0.0685***	-0.091***	-0.066***
adoptionTime	0.123***	0.1142**	0.089*	0.122**	0.080*	0.1137**	0.0800.	0.1043*	0.072.	0.097*
isDiggAdsTRUE	0.516	0.6344	1.006.	0.535	1.100*	1.6194**	1.4490*	1.7116**	1.018.	1.573**
genderf	-0.048	-0.0028	0.071	-0.029	-0.021	-0.0132	-0.021	-0.0368	-0.05	-0.087
genderm	0.379**	0.4052***	0.390**	0.384**	0.348**	0.3312*	0.3595**	0.3381*	0.364**	0.354**
rgenderf	0.184***	0.1913***	0.194***	0.186***	0.186***	0.1836***	0.1896***	0.1851***	0.186***	0.187***
rgenderm	0.109**	0.1115**	0.107**	0.110**	0.099**	0.1030**	0.1078**	0.1082**	0.101**	0.113**
logDiggs	-0.078	-0.1163	-0.146	-0.078	-0.186	-0.2327.	-0.1928	-0.2325.	-0.176	-0.235.
logComments	-0.055	-0.0343	-0.02	-0.053	-0.023	0.0061	-0.0166	0.0075	0.015	0.058
logSubmissions	-0.253**	-0.2177*	-0.193*	-0.253**	-0.139	-0.1306	-0.1206	-0.1228	-0.170.	-0.162.
logAvgDiggs	0.007	0.0412	0.101	0.007	0.172	0.2179*	0.2189*	0.2382*	0.158	0.227*
logAvgComments	0.022	-0.0007	-0.03	0.019	-0.038	-0.058	-0.0534	-0.0642	-0.076	-0.117
logAvgSubmissions	0.295**	0.2726**	0.271*	0.299**	0.200.	0.17	0.1643	0.1557	0.240*	0.207.
rLogDiggs	0.183***	0.1806***	0.165***	0.182***	0.155***	0.1504***	0.1547***	0.1521***	0.159***	0.157***
rLogComments	-0.174***	-0.1718***	-0.160***	-0.174***	-0.152***	-0.1508***	-0.1520***	-0.1520***	-0.153***	-0.154***
rLogSubmissions	-0.145***	-0.1440***	-0.132***	-0.144***	-0.128***	-0.1272***	-0.1280***	-0.1280***	-0.132***	-0.133***
rLogAvgDiggs	0.526***	0.5237***	0.527***	0.525***	0.540***	0.5418***	0.5407***	0.5411***	0.537***	0.539***
rLogAvgComments	0.282***	0.2800***	0.265***	0.282***	0.246***	0.2442***	0.2463***	0.2460***	0.248***	0.248***
rLogAvgSubmissions	0.021	0.0222	0.008	0.021	0.001	0.0007	-0.0006	0.0006	0.007	0.007
wday1	0.066	0.0452	-0.004	0.058	0.038	-0.0137	0.032	0.0083	0.019	-0.029
wday2	0.153	0.1577	0.121	0.15	0.149	0.1331	0.1725	0.1744	0.15	0.177
wday3	0.238.	0.2259.	0.260*	0.237.	0.285*	0.2609*	0.2961*	0.2892*	0.297*	0.321*
wday4	0.354**	0.3285*	0.312*	0.354**	0.268*	0.2900*	0.2567.	0.2863*	0.271*	0.301*
wday5	-0.458	-0.4275	-0.619.	-0.463	-0.666.	-0.6047	-0.7158.	-0.6409	-0.698.	-0.676
wday6	0.331.	0.2598	0.105	0.315	0.131	0.1823	0.1622	0.2012	0.084	0.151
hour(5,11]	-0.009	-0.0274	-0.114	-0.015	-0.112	-0.0881	-0.1735	-0.1313	-0.111	-0.128
hour(11,17]	0.055	0.0456	0.008	0.052	0.014	0.0193	0.0078	0.024	0.023	0.039
hour(17,23]	0.165	0.1594	0.129	0.164	0.111	0.1117	0.0843	0.0964	0.114	0.097
logLikelihood	-16865	-16863	-16850	-16865	-16838	-16825	-16828	-16821	-16835	-16816
AIC	33802	33800	33773	33804	33754	33731	33736	33725	33750	33717

Table B1. Complete Parameter Estimates from Different Model Specifications