Information Systems Research

Socially Nudged: A Quasi-Experimental Study of Friends' Social Influence in Online Product Ratings

Chong (Alex) Wang, Xiaoquan (Michael) Zhang, Il-Horn Hann

Please scroll down for article—it is on subsequent pages

Socially Nudged: A Quasi-Experimental Study of Friends' Social Influence in Online Product Ratings

Chong (Alex) Wang,[a] Xiaoquan (Michael) Zhang,[b] Il-Horn Hann[c]

[a] Guanghua School of Management, Peking University, 100871 Beijing, China; [b] CUHK Business School, Chinese University of Hong Kong, Shatin, Hong Kong; [c] Robert H. Smith School of Business, University of Maryland, College Park, Maryland 20742
Contact: alex.wang@gsm.pku.edu.cn, http://orcid.org/0000-0001-6243-7062 (C(A)W); zhang@cuhk.edu.hk,
http://orcid.org/0000-0003-0690-2331 (X(M)Z); ihann@rhsmith.umd.edu (I-HH)

Abstract. Social-networking functions are increasingly embedded in online rating systems. These functions alter the rating context in which consumer ratings are generated. In this paper, we empirically investigate online friends' social influence in online book ratings. Our quasi-experimental research design exploits the temporal sequence of social-networking events and ratings and offers a new method for identifying social influence while accounting for the homophily effect. We find that rating similarity between friends is significantly higher after the formation of the friend relationship, indicating that with social-networking functions, online rating contributors are socially nudged when giving their ratings. Exploration of contingent factors suggests that social influence is stronger for older books and for users who have smaller networks, and that relatively more recent and extremely negative ratings cast more salient influence.

1. Introduction

Online product ratings often play a useful role in informing consumers' purchasing decisions. The value of an online rating system lies in how effectively it solicits truthful expressions of private evaluations from consumers of the products, which depends crucially on the rating context in which ratings are generated.[1] In this study, we examine the generation of online ratings from the perspective of social interactions between online reviewers. We are interested in understanding how social-networking function creates a rating context, a social-choice architecture, in which users are constantly socially nudged in rating decisions (Thaler and Sunstein 2008). Leveraging on the dynamic nature of online social networks, we empirically identify a response of users' ratings to their friends' ratings. In other words, a social-networking function indeed results in online ratings being socially nudged. Although social nudge can arise from both observational learning and peer pressure, given that friends' ratings are not necessarily more accurate, deviations in users' ratings from their private information signals could undermine the usefulness of online rating systems. As more and more online rating websites integrate social-networking functions into their existing services, social nudge in online ratings is likely

to have increasingly significant consequences (Salganik et al. 2006).

This paper examines friends' social influence in online book ratings using data from a large social network-based online rating website. While theoretical discussion about social influence has been abundant, it has been empirically difficult to identify and evaluate social influence between friends using observational data (Bapna and Umyarov 2015, Manski 1993). In this study, we propose an empirical strategy, a quasi-experimental design, to identify social influence in online ratings using observational data when a confounding homophily effect is present. This methodology can be used in different online social networking contexts. Identification in the quasi-experimental design hinges on the ability to observe the time when a pair of users become friends and the time when they leave ratings for the same book before and after they become friends. By examining the similarity in ratings before and after people become friends, we can estimate the direction and magnitude of social influence in ratings. We further conduct various robustness checks to ensure the validity of the design and rule out alternative explanations, for example, endogenous timing of friendship formation.

Our results suggest that, on top of taste similarity among friends (i.e., the homophily effect), users' earlier ratings exert social influence on their friends' later ratings. On average, rating similarity between online friends is about 1.9 times higher after they become friends. Extending the research design, we examine a few contingent factors. This analysis suggests that those who have fewer friends are more easily influenced, that the influence is more salient for older books, and that more recent and extremely negative ratings cast stronger impacts.

This study contributes to the literature in several ways. First, we contribute to the online word-of-mouth (WOM) literature by studying the impact of social context on the generation of ratings. WOM studies typically assume that consumers' ratings are based on their own opinions formed after consumption (e.g., Kuksov and Xie 2010, Li and Hitt 2008, Malthouse et al. 2013). We point out that the rating context matters, and user ratings are socially nudged by their online friends' opinions. Social nudge is introduced by the implementation of social-networking features in online WOM systems. Designers and users of these systems should be aware that although social-networking features might help attract users and improve stickiness, social nudge as a result of interactions among friends may prevent users from giving independent evaluations of products.

Second, we contribute to the literature on social influence by proposing an innovative quasi-experimental design that identifies social influence between online friends on top of the homophily effect (McPherson et al. 2001). It is important to separate the effect of homophily from the effect arising out of social influence in our context, because these two effects have very different strategic implications for managers. If homophily is the dominating force behind the similarity in ratings given by friends, then managers should not be concerned about the high correlation between friends' ratings. If, instead, it is social influence that induces later reviewers to give ratings conditional on their friends' earlier ratings, then the rating-score trajectory will be path dependent, and whoever leaves the first rating will influence his friends' future ratings. Our research design exploits the temporal sequence of social-networking activities and rating events and offers an easy-to-implement way to derive causal interpretations from observational data. It explicitly takes care of endogenous friend relationship formation and the homophily effect. Unlike other empirical methods proposed to identify social influence that depend either on experimental manipulation or complex empirical assumptions, our method can be easily replicated in other contexts and scales well for big observational social-networking data sets.

Third, we contribute to the social contagion literature by empirically studying friends' social influences in postadoption opinion reporting. In previous research, social influence is usually identified on the basis of an act of consumption or adoption. Insights obtained from studies of social influence in adoption cannot be easily applied to understanding postadoption social influence in reporting, because the mechanisms through which the social influences take place in adoption and opinion reporting are likely to be different. To better understand the social influence mechanisms, we investigate contingent factors that might moderate the identified social influence. These additional results serve as both a robustness check of the proposed methodology and a starting point for managers and system designers to assess the managerial and strategic implications of social-networking features in online rating systems.

The rest of this paper is organized as follows. Section 2 introduces the background and reviews the literature. In Section 3, we discuss our research design. Section 4 introduces our empirical research context and measures. In Section 5, we present and discuss the results, robustness checks, and additional analyses, including an exploration of contingencies in social influence. Section 6 concludes.

2. Background and Literature Review

Over the past decade, we have witnessed rapid penetration of social media and social networks in various online applications. According to the Pew Research Center, by 2013, 73% of American adults were using online social-networking sites (Duggan and Smith 2013). The 2014 *U.S. Digital Consumer Report* found that 64% of social media users and 47% of smartphone owners visit social networks daily (Nielsen 2014). Attracted by the benefits of rapid viral growth and fewer fraudulent ratings, popular online rating sites are quick to embed social-networking features. Yelp (http://www.yelp.com), Rotten Tomatoes (http://www.rottentomatoes.com), and TripAdvisor (http://www.tripadvisor.com), for example, encourage users to invite friends to join the network and display friends' reviews and ratings in more prominent positions. This study examines the change in individuals' WOM reporting when they have access to their friends' ratings. Our research is broadly related to two streams of prior studies, namely, online WOM and social influence.

2.1. Studies on Online Product Ratings and Reporting Biases

Online WOM is probably the earliest form of user-generated content (UGC). Individual consumers contribute to and benefit from Internet UGC applications, such as online discussion boards (Antweiler and Frank 2004), Usenet groups (Godes and Mayzlin 2004), online exchange platforms (Resnick and Zeckhauser 2002), *Wikipedia* (Zhang and Zhu 2011, Zhang and Wang 2012, Xu and Zhang 2013), YouTube (Susarla et al. 2012,

Yoganarasimhan 2012), and online movie/game/book rating systems (e.g., Chevalier and Mayzlin 2006, Chintagunta et al. 2010, Liu 2006, Zhu and Zhang 2010). From a consumer's perspective, online ratings can significantly reduce the risk associated with the uncertainty of purchasing experience goods (Bolton et al. 2004, Pavlou and Gefen 2004). From a seller's point of view, such ratings are a valuable information channel and can be a useful marketing tool (Lu et al. 2013). Previous studies show the sales impact of various aspects of online WOM (Chevalier and Mayzlin 2006; Chintagunta et al. 2010; Dellarocas et al. 2007; Duan et al. 2008a, b; Forman et al. 2008; Godes and Mayzlin 2004; Gu et al. 2012; Liu 2006; Yin et al. 2015). As a result, firms are attentive and respond strategically to online ratings (Chen and Xie 2005, Dellarocas 2006, Hu et al. 2011, Mayzlin et al. 2014).

The value of online rating systems lies in the quality of information they deliver, which depends on the underlying mechanisms of rating generation. Dellarocas (2006) argues that although consumer ratings may still be informative when firms can manipulate online ratings, ratings generated under this mechanism can result in a social welfare loss.

A number of recent papers examine the generation of online ratings and its consequences. The literature suggests that online ratings can be biased owing to self-selection in user reporting. Li and Hitt (2008) develop a model to explain the dynamic pattern of product ratings as a result of consumers' self-selecting into early and late adopters. They empirically document that even with truthful reporting of perceived quality, early and later ratings should not be interpreted in the same way. Dellarocas et al. (2007) and Godes and Silva (2011) report a similar downward trend in product ratings. Hu et al. (2009) further identify two sources of self-selection bias, namely, acquisition bias and underreporting bias. Dellarocas and Wood (2008) find that reporting bias arises when one's propensity to report a privately observed outcome to an online reputation system depends on the type of outcome. Selective underreporting thus distorts the distribution of publicly reported ratings and renders judgments that are based solely on such ratings erroneous.

Wu and Huberman (2008) study the dynamic aspects of online opinion formation and find that exposure to previous public opinions leads reviewers into a trend-following process of posting increasingly extreme ratings. Similarly, Moe and Trusov's (2011) empirical model suggests that later ratings can be affected by earlier public ratings. Moe et al. (2011) explain that reporting bias results from a selection effect and an adjustment effect. Schlosser (2005) experimentally demonstrates a negativity bias in ratings when social concerns about self-presentation (appearing more intelligent and competent) are triggered. Marketing research had been documenting a similar influence in opinion expression

long before the existence of online ratings. Cohen and Golden (1972), for example, let subjects evaluate a brand of coffee under four different conditions with respect to information exposure and visibility expectation. They conclude that exposure to others' evaluations significantly influences subjects' ratings. In another study, Burnkrant and Cousineau (1975) find that information about prior evaluations significantly influences the ratings given by subjects. In a recent development in the literature, Goes et al. (2014) empirically demonstrate the "popularity effect" in online WOM expression resulting from user subscriptions. Burtch et al. (2017) study the impact of social norm using randomized field experiments. Huang et al. (2017) examine how social network integration affect the characteristics of online reviews.

Differing from the existing literature that focuses on self-selection, intentional distortion, and the impacts of public rating information and review subscription, we study the impacts of online friend relationships on ratings. Online friend relationships and friends' prior ratings create a social context in which users express their evaluations. This context is likely to have a significant impact on the ratings being produced. It is important to identify and acknowledge the impact of friend influence in online ratings, considering that online review systems increasingly depend on embedded social networks and people have adapted to using online social networks to maintain close social connections. In a related study, Lee et al. (2014) study the generation of online ratings from the social learning perspective and consider online friends' ratings as a source of learning. Their finding suggests that learning is present in online ratings, but public (non-friends') ratings exert greater influence than friends' ratings. Differing from their work, which focuses on observational learning, we focus on identifying friend influence in online ratings by proposing an easy-to-implement quasi-experimental research design that explicitly takes care of endogenous friend relationship formation and the homophily effect.

2.2. Studies on Social Influence

Numerous studies in social psychology demonstrate that people behave very differently when they are under social influence (e.g., Cialdini and Goldstein 2004). Research on communication networks, innovation diffusion, and opinion leadership has long recognized that consumers are influenced by others (e.g., Van den Bulte and Lilien 2001). We focus on social influence from online friends. Information about friends' ratings casts influence on focal users' rating behavior through two general mechanisms, informational influences and normative influences (Burnkrant and Cousineau 1975, Cialdini and Goldstein 2004, Deutsch and Gerard 1955), which researchers also refer to as observational learning and peer pressure,

respectively (e.g., Cai et al. 2009, Liu et al. 2015, Mas and Moretti 2009, Moretti 2011, Zhang 2010). Through observational learning, friends' ratings convey new information about the product being reviewed that a user can rely on to update his evaluation. Peer pressure, by contrast, refers to a user's tendency to conform to friends' ratings motivated by positive identification with friends and the intention to maintain close social connections. Postconsumption evaluation typically involves little uncertainty, and friend influence results mainly from peer pressure. Some products or services (e.g., dietary supplements, exotic restaurants, and expert services, such as medical procedures and automobile repairs), which are often referred to as credence goods, however, have the feature that consumers may have difficulty evaluating their quality even after consumption (see Dulleck and Kerschbamer 2006). For these goods, a focal user's rating may be influenced when he attempts to infer/learn the goods' quality from his friends' ratings in addition to the social pressure. As we will discuss in Section 5, our results from online book ratings favor an explanation based on peer pressure.

Much of the literature on social influence examines product and innovation *adoption* under uncertainty (e.g., Cai et al. 2009, Zhang and Liu 2012). Differing from these studies that focus on adoption, our paper examines opinion reporting. When consumers face preadoption uncertainty in products, herding (following others' actions without utilizing their own private information) can be a viable equilibrium strategy that results from observational learning (Banerjee 1992, Bikhchandani et al. 1992). Insights obtained from studies of social influence in adoption, however, cannot be easily generalized to understand friends' social influence in opinion reporting, because the mechanisms through which social influence takes place in adoption and opinion reporting are likely to be different.

Identifying friends' social influence in ratings is challenging because one cannot simply use the strong correlation in friends' ratings as evidence of their influencing each other: Strong correlation in ratings can also result from similarity in friends' tastes (the homophily effect) or, equivalently, the endogenous formation of friend relationships (Lazarsfeld and Merton 1954, McPherson et al. 2001). Homophily refers to the phenomenon that socially proximate individuals tend to have similar individual-level characteristics. Thus, similarities in their behavior may be driven by common characteristics that are often unobserved. Distinguishing social influence from homophily and other confounding factors is a well-known empirical challenge (e.g., Manski 1993).

Various solutions have been proposed for distinguishing the effect of social influence from that of other relevant factors (Brock and Durlauf 2001, Soetevent 2006). The ideal method would entail conducting randomized experiments, by assigning individuals into groups with different treatment conditions and then examining the effect of social influence. Such randomized experiments are typically very costly to conduct because it is difficult to manipulate social ties. Field and quasi experiments are valid alternatives. Sacerdote (2001) examines peer influence on academic performance with a randomized sample of college students (see also Foster 2006). In their study of productivity spillover, Mas and Moretti (2009) leverage the quasi-random arrangement of working shifts. In a study of retirement plan enrollment, Duflo and Saez (2003) randomly vary the level of social interactions among potential participants and infer the impact of social interaction from the identified spillover effects. A limitation of these studies is that social interactions and social ties in the research context are often not directly observed and measured. In our study, the complete history of the social network's development is recorded, making it easier for us to measure the social relationships among users. Researchers also have conducted online randomized field experiments to identify social influence between online friends in product adoption (Bapna and Umyarov 2015). These studies offer strong evidence of social influence between online friends, but randomized experiments are costly to replicate. The method proposed in this paper is based solely on observed social connections that are readily available to site managers.

A second approach relies in econometric manipulations, such as adding fixed effects and explicitly modeling the selection process. Identification can leverage the panel-data structure of social influence over time (Brock and Durlauf 2001) or the structure of network interactions (Bramoullé et al. 2009). A stochastic actor-based modeling approach was proposed recently to model the coevolution of social networks and behavior (Lewis 2011, Snijders et al. 2006, Steglich et al. 2010). To compensate for the lack of empirical control and observations, these models tend to have strict requirements for the identification conditions (e.g., Angrist and Pischke 2010, Bollen and Pearl 2013, Summers 1991). By contrast, our identification does not require such strong modeling assumptions.

Social-interaction effects can also be estimated by exploiting natural instrumental variables or exogenous shocks (e.g., Brown et al. 2008, Conley and Udry 2010, Tucker 2008). Researchers engaging in this type of research leverage the richness of data to find creative ways of identification.

In our quasi-experimental design, we exploit the ratings' visibility and the dynamic feature of social networks to eliminate the homophily effect and identify social influence in friends' ratings.[2] Our empirical results confirm the existence of the homophily

effect and reveal that the generation of online ratings is subject to friends' influence. Our approach is easy to implement in alternative social-network environments, especially in online social networks that feature large social groups and voluminous user activities. It offers a way to examine large-scale social interactions in contexts where implementing a full-scale randomized experimental design is infeasible.

3. Research Design
3.1. Identification of Social Influence: A Quasi-Experimental Design

Our estimation strategy builds on a response function of focal user i's rating for book j, or $Rating_{ij}$, on i's friends' average rating of the same book, $AvgFrdRating_{ij}$, controlling for other user-book-specific factors at the time of the focal user's rating, X_{ij}

$$Rating_{ij} = f(AvgFrdRating_{ij}, X_{ij}).$$

For $AvgFrdRating_{ij}$, we consider only the ratings for book j left by the friends of user i before the focal user i's rating (for book j). An important concern here is that since friend relationships are formed endogenously, the correlation between a focal user's rating and his friends' average rating may not be the result of social influence, but a consequence of their sharing similar tastes. Similarity in tastes, or the homophily effect, confounds the social-influence interpretation of the response function. To tease out social influence from the homophily effect, an ideal experimental environment would require picking subjects randomly, manipulating the visibility of friends' previous ratings, and then examining whether the subject's action differs under various visibility treatment conditions. Such experiments, however, are hard to conduct on a large

scale in functioning social networks. To achieve similar rigor in identification while taking advantage of naturally available observational data, we can rely on a quasi-experimental design (Campbell and Stanley 1963). As we show below, with certain testable assumptions, quasi randomization over rating visibility can be achieved based on the timing of both ratings and friend relationship formation.

Figure 1 depicts the (relative) timing of three events that we leverage in the empirical framework. "FocalR" indicates the event when the focal user gives a rating. "UserAR" indicates the event when a friend of the focal user (user A) leaves a rating in the system. Finally, "A&F" denotes the event when the two users become friends.

Panel A shows the AFTER case, in which the focal user's rating (FocalR) takes place *after* the friend relationship forms (A&F). In this case, the relative timing between UserAR and A&F is not critical, since it is only required that UserAR takes place before FocalR. Panel B shows the BEFORE case, in which the focal user leaves the rating (FocalR) before he becomes friends with user A (A&F).

As a result of the existing online friend connection, user A's rating is salient in the AFTER case in panel A, but not in the BEFORE case in panel B. If the similarity between the two users' ratings is stronger in the AFTER case, we would then have evidence to support that the rating has been influenced by the social connection.[3] In other words, to separate social influence from the confounding homophily effect, we examine the similarity in ratings before and after the online friendship occurs, using the BEFORE case as the benchmark of inherent rating similarity between friends.

Each focal user's friends are defined according to their relationship by the end of the observation period (the complete friend network). For friends of a focal user

Figure 1. Illustration of the Relative Timing of Ratings and Social-Networking Events
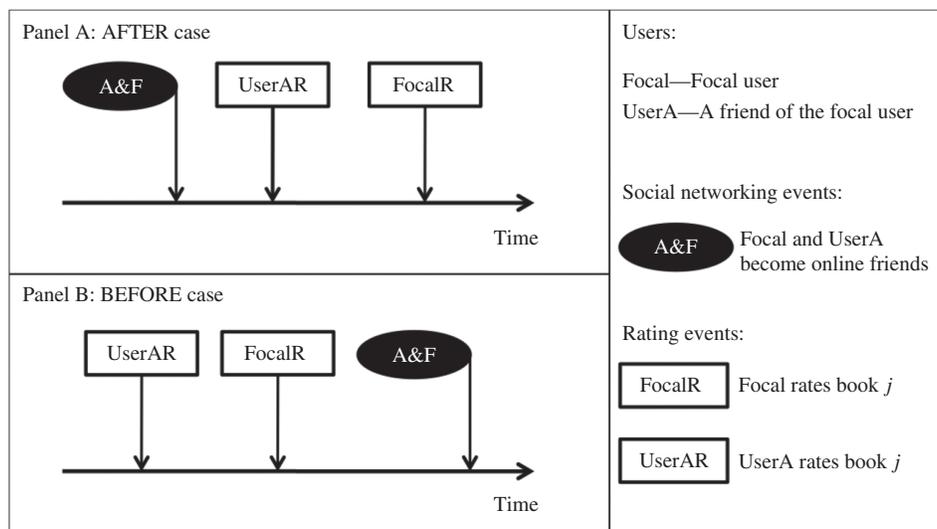
who have rated the same book, we define a dummy variable to indicate whether the focal user's rating takes place before they become friends ($After_{ij} = 0$, panel B in Figure 1) or after they become friends ($After_{ij} = 1$, panel A in Figure 1). The variable $After_{ij}$ therefore indicates whether, at the time of focal user i's rating of book j, the rating of the same book by his friend (that happened before i's rating of book j) is salient to the focal user as a friend's rating or not. Since visibility of an influencers' behavior is the single most important precondition for social influence to take place (Marsden and Friedkin 1993, Mas and Moretti 2009), we examine the parameter estimate of the interaction between $After_{ij}$ and $AvgFrdRating_{ij}$ to identify the social influence. Without social influence, the similarity in friends' tastes should remain the same regardless of whether the focal user can view his friends' ratings. We then would expect to see no effect of $After_{ij}$ on the rating similarity (the relationship between $Rating_{ij}$ and $AvgFrdRating_{ij}$). If there is social influence, we should identify a significant interaction effect between $After_{ij}$ and $AvgFrdRating_{ij}$.

Based on our research design, the traditional linear-in-mean social interaction model (Brock and Durlauf 2001) that allows for variations across books and users with other control variables can be written as follows:

$$
\begin{aligned}
Rating_{ij} = \alpha &+ \beta_1 AvgFrdRating_{ij} + \beta_2 After_{ij} \\
&+ \beta_3 AvgFrdRating_{ij} \times After_{ij} + X_{ij}\gamma \\
&+ u_i + v_j + \varepsilon_{ij}.
\end{aligned} \tag{1}
$$

Since individuals other than friends of the focal user may have rated the same book,[4] to ensure that both $After_{ij}$ and $AvgFrdRating_{ij}$ can be calculated in the linear-in-mean model, we need to require that these friend ratings all belong to either the AFTER case or the BEFORE case, but not both. It is possible that some of the friends' ratings belong to the AFTER case and others belong to the BEFORE case, and thus we examine these cases separately based on a similar research design to offer corroborating evidence in Section 5.2.

3.2. Discussion of the Identification Strategy

As our research design is based on observational data, several things cannot be controlled. First, the pairs of friends cannot be randomized (endogenous friend relationship); second, the time when two users become friends cannot be manipulated (endogenous timing of friend relationship formation); and finally, the order of giving ratings is self-selected (endogenous timing of ratings). We next explain how these concerns are addressed in this study.

Endogenous Friendship—Homophily. Endogenous friendship, or homophily, refers to the fact that people select their friends based on common interests. In observational studies, friendship formation is endogenous. As we reviewed, studies of social influence also

often rely on observed endogenously formed social ties. Duflo and Saez (2003) and Tucker (2008), for example, study social influence carried by endogenously formed coworker friend circles. Brown et al. (2008) examine social influence among naturally formed neighbors. One important objective of these studies is to propose methods to tease social influence out from homophily. We similarly address naturally formed online friend ties in the research design.

It is important to point out that our research focuses on pairs of users who eventually become online friends. The research design involves no comparison between friends and strangers. Our research design leverages on the relative temporal order of friend relationship formation and ratings to create treatment and control groups. In other words, we are comparing friends at different time points. As with all quasi-experimental designs, it is crucial to assess whether the treatment can be considered reasonably random. Consequently, it is important to discuss the validity of the quasi-experimental design in terms of the randomness of the timing of these events.

Endogenous Timing of Friend Relationship Formation. Endogenous timing of friend relationship formation refers to the problem that people not only self-select to be friends with certain people, but they may also self-select the time when they become friends with others. This is the most significant challenge to our design. A crucial assumption to be satisfied is that the temporal sequence of users becoming online friends with each other is not systematically related to the similarity between them. If earlier friend relationships indeed exhibit higher similarity than later ones, because of the cumulative nature of online ratings, a temporal sampling bias that is common to quasi-experimental designs would arise. In this case, in calculating $AvgFriendRating_{ij}$, we would sample more shared ratings from similar friends than from dissimilar friends in the AFTER case, because similar friends tend to form friend relationships earlier than dissimilar friends. If this happened, social influence identified by the $After_{ij} \times AvgFriendRating_{ij}$ would overestimate the actual social influence.

To address this concern, we need to rule out the possibility that earlier friends are more similar than later friends. We carry out a few robustness checks. First, we take into consideration the tenure of the friend relationships between users. We demonstrate with two analyses that earlier friends are no more similar to focal users than later friends. Second, we consider an additional analysis at the friend-pair level and examine the rating similarity between the same pair of friends. Since this estimation is on the friend-pair level, the timing of friend relationship formation could be further controlled by friend-pair fixed effects. Finally, we demonstrate that a pair of friends does not naturally

become more similar over time. Based on the dyad-level analysis, we show that (1) people in a friend pair do not become more similar before the introduction of the social-networking function, and (2) introduction of the social-networking function alone does not trigger higher similarity in friend pairs: only when two users become friends (and thus can view each other's ratings) does social influence take place. These additional analyses are reported in Section 5.2.

Endogenous Timing of Ratings. In a perfectly randomized experiment, subjects' roles are selected before the experiment. Subjects in the treatment group will see their friends' earlier ratings and those in the control group will not see their friends' ratings. In our design, we cannot pick subjects' roles up front. Focal users' ratings (in both the control group and the treatment group) are always the later (relative to friends') ratings by the users.

This self-selected rating order does not affect the validity of our design and results. First, our empirical test hinges on whether friends' ratings are visible or not. Even if later ratings are systematically more similar to or different from earlier ratings, we should not find any significant difference before and after when friends' ratings are visible unless there is social influence. Second, it is possible for some users to change their habits after implementing the friend function so that they wait longer to see their friends' ratings before giving their own ratings, but this is in line with our proposition that online ratings are socially nudged with the friend network function. In our robustness checks, we examine whether the friendship function alters users' responses to friends' ratings.

Social Influence Before Friend Relationships. On social-networking sites, users may keep track of ratings by other users before forming online friend relationships, because the formation of such relationships requires mutual recognition. As a result, users are influenced by their friends even before friend relationships are formed. Although our research design leverages on the observation of friend relationship formation, we make no assumption that social influence between online friends exists only after the friend relationships are formed. If there is social influence even before friend relationship formation, our estimation will underestimate the actual influence.

4. Research Context, Data, and Measures
4.1. Research Context
To implement the above research design and test the significance of social influence in online ratings, we obtained data from one of the most influential online rating websites for books, movies, and music in China. Established in 2005, the site has more than 8 million registered users and attracts more than 10 million page views per day. These page views can be from either registered or unregistered users. Registered users can leave ratings and write reviews about items that they have consumed and gradually form an online profile that serves as the foundation of social-networking activities on the site, whereas unregistered users mainly browse the site to acquire information about books, movies, and music.[5]

Through the search and browse functions, users can rate items and form a personal collection of books, movies, and music. The site's collaborative-filtering algorithm uses information from user collections to suggest new items and potential social connections. When reviewing an item, a user can express his opinion and choose a star rating from one to five. All ratings and reviews are public.

The site introduced a friend network function in February 2008. With the friend network function, friends' activities are shown conspicuously. Friend requests can be easily initiated with a click of a button on users' profile pages. If a targeted user agrees to a friend request, an online friend relationship will be formed. Once two users form a friend relationship, each will be updated about the other's activities, including ratings. The social-update mechanism makes the individual friends' ratings distinctively salient and separate from the ratings by other users, which are presented in aggregate form on the book page. (Detailed descriptions of the sites can be found in the online appendix.) Salient information about friends' ratings is a critical condition for social influence to take place. On one hand, the salience of a friend's rating-information feed enables a user to be aware of his friends' expressed opinions. On the other hand, users are also aware that their friends will be able to easily access their expressed opinions.

The site promotes friend relationships by collecting user preference data and provides users with information about their common interests with other users. If a user visits another user's profile page, the site will automatically show the items that they both liked. While being the most influential user-review site for cultural products in China, it is strategically positioned as a social-networking site and does not feature functions that recognize users' contributions as reviewers as much as other review sites (e.g., Yelp gives badges to differentiate reviewers). Only users' collections and activities are shown on user profile pages. There is no salient information that vertically differentiates users. Decisions to initiate friend relationships are based mostly on common interests and on-site social interactions. The site also provides a messaging function that enables users to communicate.

4.2. Data and Measures
We collected the data from the site's data server archive, which contains the entire history of users' ratings for items (including books, movies, and music)

from 2005 to 2008. The complete data set has about 50 million ratings for over a half million items from about 890,000 users. We also observe the social network typology. The entire social network in our data set contains over 2 million links among 286,140 users. Since the friend network function was introduced in February 2008, in the main analysis, we focus on observations of book ratings from February to August 2008. In the robustness tests, we leverage the availability of data beyond this time window.

As discussed in Section 3, our empirical identification relies on the relative timing of friend relationship formation and friend ratings. In the BEFORE case, $AvgFrdRating_{ij}$ is given before friend relationships form, while in the AFTER case, $AvgFrdRating_{ij}$ is given after friend relationships form. In Equation (1), $After_{ij}$ is a dummy variable that indicates whether the $AvgFrdRating_{ij}$ belongs to the BEFORE case ($After_{ij} = 0$) or the AFTER case ($After_{ij} = 1$). The interaction between $After_{ij}$ and $AvgFrdRating_{ij}$ (β_3) thus identifies the social influence. To clearly define $After_{ij}$, we require that friends' ratings for calculating $AvgFrdRating_{ij}$ are either all from the BEFORE cases or all from the AFTER cases. This process gives us a data set of 171,588 ratings, covering 20,480 book titles by 33,605 users.[6]

Control Variables. In Equation (1), we include various measures of rating, book, and user characteristics as controls. We capture the decay of social influence with a variable that measures the number of days from friends' last rating to the time of the focal rating, $Recency_{ij}$. A lower value of $Recency_{ij}$ indicates that friends' ratings are more recent.

In terms of book characteristics, we calculate book age ($BookAge_{ij}$), measured by the number of days from the time book j appeared in the data set to the time of the focal rating, and rating intensity ($RatingIntensity_{ij}$), measured by the average number of ratings per day before the focal user's rating. To control for general opinions on each book, we also include the count, average, and variance of the ratings for book j of all users at the time of the focal rating ($NumRating_{ij}$, $AvgRating_{ij}$, and $VarRating_{ij}$). An average book in our data set gets a rating of 4.1 on a five-star scale. Before getting each focal rating, an average book has been on the site for about 829 days since its first rating and has received 2,894 user ratings. In addition to these covariates, we also introduce book fixed effects to control for how book characteristics may affect the similarity between the focal rating and the focal user's friends' ratings.

As for the users, we control user experience, measured by the number of days from user i's first appearance in the data set to the time of the focal rating ($UserAge_{ij}$), the number of friends that user i has ($NumFrd_{ij}$), and the number of books that user i has rated ($NumBook_{ij}$) by the time of the rating. Definitions

and summary statistics of variables are summarized in Table 1.[7] On average, users in our data set have 17 friends. Before the focal rating, on average, a user has been using the system for 244 days and has rated 164 books. While users are quite active in rating, the number of friends who rated the same book before a focal user (number of friend ratings) is not high, averaging 1.418 in the AFTER cases and 1.314 in the BEFORE cases. Less than 10% of the focal users' ratings have more than three prior friends' ratings.

5. Results and Discussion
5.1. Estimation Results
Estimation results for the linear-in-mean model are reported in Table 2. In addition to observable user and book characteristics, we control for user and book fixed effects in both models. In column (1), we first report estimates from a "naïve" model, in which we estimate the correlation between the focal rating and friends' previous ratings without considering the relative timing of focal user's ratings and the formation of friend relationships. As expected, $AvgFrdRating_{ij}$ is significant and positive, indicating that friends' ratings are similar to each other. The similarity in focal users' ratings and their friends' ratings, however, may be a result of both the homophily effect and the social influence effect. Consistent with previous studies on online WOM dynamics (Moe and Schweidel 2012), a focal user's rating is lower when (a) he is more experienced ($NumBook_{ij}$) and (b) the book has been more intensely rated ($RatingIntensity_{ij}$).

Column (2) of Table 2 reports estimates from our main model, in which we introduce the "treatment" variable, $After_{ij}$, and its interaction term with $AvgFrdRating_{ij}$. The positive and significant interaction term suggests that social influence from friends' ratings indeed exists. A back-of-the-envelope calculation suggests that, on average, rating similarity almost triples (increases by 190%) after users become friends.[8] The coefficient of $AvgFrdRating_{ij}$ is positive and significant, indicating that over 30% of the rating similarity identified in the naïve model comes from the homophily effect. Coefficients for the group difference control variable ($After_{ij}$) are marginally significant and negative, indicating that the focal users' ratings in the AFTER cases are generally lower than in the BEFORE cases. After controlling for the book-fixed effect, the average of previous public ratings ($AvgRating_{ij}$) appears to be negatively correlated with the focal user's rating. This is consistent with existing literature indicating that individual raters exhibit a tendency of diverging from public ratings (Moe and Trusov 2011). Given the variances of the two variables, our estimation results suggest that public ratings and friends' ratings have distinctive yet equally strong impacts on focal users' ratings.

Table 1. Variable Definitions and Summary Statistics

Variable name	Definition	Summary statistics: Mean (Std. dev.)	
		Original	Logged
Rating variables			
$Rating_{ij}$	Focal user i's rating for book j	4.111 (0.845)	1.388 (0.242)
$AvgFrdRating_{ij}$	Average rating for book j given by the focal user's friends before the focal rating	4.134 (0.810)	1.396 (0.233)
$After_{ij}$	A dummy variable that equals 1 in cases where $AvgFrdRating_{ij}$ is from users who had become friends of the focal user (AFTER cases) and 0 otherwise (BEFORE cases)	0.545 (0.498)	
$Recency_{ij}$	Days from friends' last rating to the time of the focal rating	191.1 (214.3)	4.426 (1.580)
Book variables			
$BookAge_{ij}$	Days from book j's first appearance in the data set to the time of the focal rating	828.8 (342.1)	6.545 (0.763)
$RatingIntensity_{ij}$	Average number of ratings per day for book j before the focal rating	8.828 (15.30)	1.608 (1.150)
$AvgRating_{ij}$	Average rating (valence) for book j given by other users before the focal rating	4.088 (0.343)	1.404 (0.0878)
$NumRating_{ij}$	Volume of user ratings for book j before the focal rating	2,894.4 (4,339.7)	6.548 (2.126)
$VarRating_{ij}$	Variance of user ratings for book j before the focal rating	0.611 (0.205)	0.469 (0.125)
User variables			
$NumFrd_{ij}$	Number of friends that focal user i has made before the focal rating	17.19 (38.46)	2.000 (1.294)
$UserAge_{ij}$	Days from focal user i's first appearance in the data set to the time of the focal rating	243.8 (272.6)	4.173 (2.258)
$NumBook_{ij}$	Number of ratings focal user i gave to other books before the focal rating	163.5 (423.6)	4.055 (1.510)
Number of users		33,605	
Number of books		20,480	
Number of obs.		171,588	

5.2. Robustness Checks and Additional Analyses

As discussed in Section 3.2, we conduct several robustness checks and additional analyses to offer corroborating support.[9]

Endogenous Timing of Friend Relationship Formation. As discussed in Section 3.2, our research design relies on an assumption that the temporal sequence of two users becoming online friends is not related to the similarity between them. A natural concern is that a user's earlier friends might be more similar to him than his later friends. If this is the case, the significant and positive interaction term in the main model may be attributable to a temporal sampling bias generated by the difference in similarity of friends added at different times (i.e., friend ratings in the AFTER cases ($After_{ij} = 1$) are more likely to be from those earlier friends who are more similar to the focal user).

To test whether this potential difference in similarity between earlier and later friends is a serious concern, we explicitly consider the effect of the tenure of friend relationship on rating similarity. The rationale is that if earlier friends are more similar to a focal user, we should observe that the focal user's ratings are more similar to the ratings of his older friends. Our tests reveal no evidence that older friends are more similar than newer ones. In other words, endogenous timing of friendship formation is not a serious concern in our data set. Actually, if we consider cases in which at least two of the focal users' friends had given ratings to the book before the focal rating, there is marginally higher similarity between the ratings of the focal users and their newer friends. This is consistent with the intuition that newer friends may actually get more attention from the focal user and thus exert higher influence.

Dyad-Level Analysis. Our main analysis is a linear-in-mean model based on examining the similarity between a focal user's rating and the average rating of his friends (Brock and Durlauf 2001). Aggregating friends' ratings is desirable in the sense that the measure captures all friends' opinions. The model is

Table 2. Friends' Social Influence in Online Product Ratings

	(1)	(2)
	DV: $Rating_{ij}$	DV: $Rating_{ij}$
$AvgFrdRating_{ij}$	5.09e–02***	2.59e–02***
	(2.87e–03)	(3.88e–03)
$After_{ij}$		–3.13e–03*
		(1.80e–03)
$AvgFrdRating_{ij} \times After_{ij}$		4.94e–02***
		(5.16e–03)
Controls		
$AvgRating_{ij}$	–7.45e–01***	–7.49e–01***
	(6.90e–02)	(6.90e–02)
$NumRating_{ij}$	–1.62e–02*	–1.62e–02*
	(9.39e–03)	(9.39e–03)
$VarRating_{ij}$	8.63e–02***	8.45e–02***
	(2.62e–02)	(2.62e–03)
$Recency_{ij}$	9.71e–04**	9.14e–04**
	(4.55e–04)	(4.65e–04)
$BookAge_{ij}$	–1.01e–02	–9.41e–03
	(1.07e–02)	(1.06e–02)
$RatingIntensity_{ij}$	–3.10e–02**	–3.11e–02**
	(1.30e–02)	(1.30e–02)
$NumFrd_{ij}$	8.17e–05	9.55e–04
	(1.29e–03)	(1.38e–03)
$UserAge_{ij}$	–1.35e–03	–1.28e–03
	(1.06e–03)	(1.06e–03)
$NumBook_{ij}$	–1.65e–02***	–1.66e–02***
	(1.15e–03)	(1.15e–03)
User fixed effects	Yes	Yes
Book fixed effects	Yes	Yes
Number of users	33,605	33,605
Number of books	20,480	20,480
Number of obs.	171,588	171,588
Adjusted R^2	0.336	0.337

Notes. Model 1 (the naïve model) is a linear-in-mean model controlling user and book fixed effects. Model 2 (the main model) identifies social influence based on our empirical strategy controlling user and book fixed effects. We use the dummy variable approach to control for fixed effects in the model. All continuous variables are log transformed and centralized. Regression results based on original data values are qualitatively the same. Standard errors are reported in parentheses.

Significance levels are displayed as *$p < 0.1$; **$p < 0.05$; and ***$p < 0.01$.

vulnerable to the endogenous timing of friendship formation. To further alleviate the concern about unobserved dyad-level heterogeneity, following the same research design, we compare the rating similarity of books between the same pair of users before and after they become friends (dyad-level analysis) to offer corroborating evidence. The dyad-level model is able to fully control dyad-level unobserved similarity and thus is robust to endogenous timing of friendship formation. Our estimation results confirm the existence of social influence.

In addition to confirming the existence of social influence using a different level of analysis, the dyad-level analysis allows us to explore (1) whether rating pairs given in a shorter time window exhibit stronger influence and (2) whether rating pairs given after the friend function introduction are systematically different from the pairs given before the function.

Our findings suggest that the magnitude of social influence is significantly larger in a subsample consisting of shared ratings given within a 10-day window. Since reading a book requires time, this result suggests that social nudge is more significant in the postconsumption stage; that is, it is more likely that focal users read friends' ratings after reading the book (postconsumption influence) rather than read friends' ratings before reading the book (preconsumption influence).[10] Our findings also suggest that rating pairs given after the friend function introduction are not systematically different from the pairs given before the function. In other words, merely introducing a friend function has no significant impact on rating similarity. Social influence takes place only after the formation of friend relationships.

Do Public Ratings Have the Same Conformity Pressure? Suppose that users tend to agree more with each other's opinions over time. Even without social influence from making friends, one might still observe ratings becoming more similar. To assess the possibility of unobserved systematic changes in how users respond to previous ratings, we examine whether the average of public ratings, $AvgRating_{ij}$, has different impacts on the focal users' ratings in the BEFORE and AFTER cases. Controlling for the friends' ratings, we find no significant treatment effect on public ratings; that is, the increase in rating similarity takes place only with friend ratings, even when we introduce public ratings into the model. This result supports our finding that the identified social influence is indeed due to conformity pressure among friends.

Alternative Empirical Model Specifications. To further assure the robustness of our findings, based on our research design, we consider a few alternative model specifications.[11] Specifically, we considered (1) using ordered logit models to account for the discrete rating scale, (2) using the deviations of focal users' ratings from public ratings as the dependent variable and deviations of friends' ratings from public ratings as the independent variable, and (3) directly comparing the absolute differences between focal users' ratings and their friends' ratings in the BEFORE and AFTER cases. The estimation results suggest that the findings from our main model are robust to alternative model specifications.

Additional Analysis of Rating Similarity in the BEFORE and AFTER Periods. As explained in Section 3.1, in the main analysis, we only consider cases where friends' ratings are either all from the AFTER period or all from the BEFORE period. Cases in which multiple friends'

ratings belong to both the AFTER and BEFORE cases are analyzed separately. Specifically, for this sample, we examine whether a focal user's rating is more similar to the ratings that belong to the BEFORE cases or the AFTER cases. As detailed in the online appendix, paired mean comparisons support the existence of social influence.

5.3. Contingent Social Influence

The literature on social influence suggests that the magnitude of social influence can be contingent on other factors. In this section, we extend our main model to examine contingency factors. The exploration of potential contingent factors offers many utilities to the current study. First, we would like to demonstrate that our methodology is well suited for studying such moderators in similar contexts. When contingent factors are different or when additional covariates are available (e.g., user demographics, product characteristics, etc.) in similar situations, the empirical strategy can be easily adapted to examine another set of moderators. Second, an examination of moderators

of social influence in online product ratings is interesting and important in and of itself (Godes 2011). Answering the why, when, and how questions of social influence holds promise as a means of deepening our understanding of the underlying mechanisms through which social influence takes place and offers the potential of providing practical guidance for marketing managers and system designers to improve their use of social-networking features. Third, contingencies revealed from the analysis suggest that the identified social influence changes opinions expression rather than induces a shift in user taste.

Estimation results reported in columns (1) and (2) of Table 3 investigate the moderating role of the valence of friend ratings. In the online-ratings context, extreme ratings convey strong feelings about a product and have more significant impacts on others. Based on the summary statistics reported in Table 1, which reveal that online ratings are generally positive, we categorize an average friend rating (*AvgFrdRating*) as extremely positive if it is higher than four stars and extremely negative if it is lower than three stars. We then replicate

Table 3. Contingent Social Influence

	(1) DV: $Rating_{ij}$ ($AvgFrdRating_{ij} \leq 3$)	(2) DV: $Rating_{ij}$ ($AvgFrdRating_{ij} > 4$)	(3) DV: $Rating_{ij}$
$AvgFrdRating_{ij}$	1.48e−02	−2.01e−02	2.41e−02***
	(1.31e−02)	(3.13e−02)	(4.49e−03)
$After_{ij}$	2.71e−03	1.88e−03	−5.02e−03***
	(9.85e−03)	(9.20e−03)	(1.90e−03)
$AvgFrdRating_{ij} \times After_{ij}$	6.23e−02***	1.81e−02	5.98e−02***
	(1.89e−02)	(4.35e−02)	(5.87e−03)
Contingent factors			
$Recency_{ij}$			−1.13e−02***
			(3.54e−03)
$BookAge_{ij}$			2.67e−02***
			(7.85e−03)
$RatingIntensity_{ij}$			3.58e−03
			(4.96e−03)
$NumFrd_{ij}$			−1.94e−02***
			(5.41e−03)
$UserAge_{ij}$			3.08e−03
			(2.83e−03)
Control variables	Yes	Yes	Yes
User fixed effects	Yes	Yes	Yes
Book fixed effects	Yes	Yes	Yes
Number of users	30,578	65,609	33,605
Number of books	21,596	49,519	20,480
Number of obs.	33,038	70,930	171,588
Adjusted R^2	0.356	0.303	0.337

Notes. This table reports estimation results of moderating effects. In column (1), the main model is replicated on a subsample where the $AvgFrdRating_{ij} \leq 3$ (extremely negative). In column (2), we consider another subsample where the $AvgFrdRating_{ij} > 4$ (extremely positive). In column (3), we include additional contingent factors in the model. In column (3), moderating effects are tested with three-way interaction terms. For $Recency_{ij}$, for example, the estimate reported in the "contingent factors" part corresponds to the regression coefficient for $Recency_{ij} \times AvgFrdRating_{ij} \times After_{ij}$. All second-order interactions are included in the model (Irwin and McClelland 2001). Standard errors are reported in parentheses.

Significance levels are displayed as ***$p < 0.01$.

our main model on two subsamples: a sample with extremely negative friends' ratings (column (1)) and a sample with extremely positive friends' ratings (column (2)). Our results suggest that social influence is more salient for extremely negative ratings, while there is no such evidence in the subsample with extremely positive ratings.

In column (3) of Table 3, we further include contingency factors that capture the characteristics of friends' ratings, books being rated, and focal users in our data set. We selected these variables with guidance from relevant discussions in the social influence literature and constraints due to data availability.[12]

First, we find that more recent friends' ratings indeed have a higher influence on focal ratings. This result suggests that friends' social influence in online WOM tapers off over time. As time passes, previous friends' ratings can become less relevant to focal users. Second, friends' influence is more salient for older books. Yet, there is no significant relationship between rating intensity and social influence. This suggests that, rather than following the mainstream, users conform to their friends in an attempt to develop and manage relationships that they regard as defining themselves in the community. This finding also suggests that while users could learn from others' ratings in the rating generation, especially when postadoption evaluation uncertainty is high, it might not be the dominant social influence mechanism through which social nudge in online ratings works. Third, having more friends implies a reduction in the average salience of friends' social influence to a user. This indicates that exposure to more friends and more friends' ratings dilute the influence. This finding is consistent with evidence in the recent literature on online social networks, which suggests that it is harder to attend to all friends as friend networks become larger (Trusov et al. 2010, Watts and Dodds 2007). Finally, we do not find a significant moderating effect of user experience as measured by $UserAge_{ij}$. This suggests that experienced users are subject to social influence to a similar extent as new users.

6. Conclusion

Using book ratings and online social-network data from a popular online rating website in China, we investigate friends' peer influence in online ratings. Our methodology exploits the temporal sequence of the formation of online friend relationships and rating activities and offers a quasi-experimental methodology to identify the presence of friends' social influence in the generation of online ratings. We examine the validity of our research design with numerous tests to extend our understanding about friends' influence in online social networks. We find that social influence is stronger for more popular books and for users with relatively smaller friend networks. In addition, extremely negative and more recent friends' ratings tend to exert greater influence.

Our results offer important managerial implications to marketers and online rating-system designers. Systems designers, depending on their objectives, can use our results to nudge their users (e.g., create or avoid social influence in opinions by adopting new functions or changing existing ones to alter the rating environment's social context). For example, rating sites can develop algorithms to recommend reviews not subject to the influence of social ties, highlight only reviews from users who do not have friends posting before them, or post a warning sign whenever it is suspected that a review might be influenced by friends, etc. For marketing practitioners, it is important to identify early adopters and take their social influence into consideration when making plans to respond to online consumer ratings and reviews. We argue that WOM management in social networks can be very different from the situation where ratings are given independently. Our additional analysis of the moderating effects of book and user characteristics can help managers effectively target their efforts to achieve marketing goals. For example, managers are likely to expend more resources on products that receive intensive user reviews, but our analysis shows that peer influence is also greater in this case, which may potentially undermine these marketing efforts. The fact that peer influence is stronger for older books suggests that for products with long life cycles, peer influence in ratings should be carefully considered. Our finding that social influence is stronger for users with small social networks, which that the issue of social influence in online ratings is particularly problematic in online ratings systems with newly introduced social networks.

Our paper makes several contributions to the literature. First, we propose a method to assess the level of friends' social influence in online product ratings, after eliminating the homophily effect that often confounds the identification of social effects. Our approach can be easily replicated in other online rating systems with social-networking features and does not require changes in the systems' functionality. The method makes it possible to evaluate peer influence in user-generated content production when only historical and observational data are available and when randomized experiments are hard to design or deploy. Compared with other methods proposed in the literature to identify friends' social influence, the quasi-experimental design also has the positive features of being computationally less demanding and theoretically less constrained. Our empirical analysis demonstrates the power of this method in dealing with big data sets with millions of ratings and social-network ties.

Second, this study differs from previous studies of social influence in two important ways: (1) While previous studies examine social influence in adoption, we study opinion reporting. The underlying mechanisms through which social influences take place can be markedly different in these two approaches. (2) While previous studies examine the public's social influence in the generation of online product ratings, we specifically show that friends exert disproportionately greater influence than the public and that the direction of influence can be different.

Third, as shown by our exploration of moderators, our research design can be easily adapted to consider contingencies in social influences. The current exploration not only offers managerial implications for managers and system designers to develop better online rating systems but also opens the door for future theoretical investigations of the underlying processes of friends' social influence on opinions.

Last, we are among the first to document friends' influence in online ratings arising from social-networking functions that are common among UGC sites. Although social networks are generally valuable in enabling efficient communication of information as well as motivating participation, social influence in opinion expression may render online ratings less useful in conveying new information. Compared with other behavioral tendencies in online ratings reported in the literature, the impact of friend influence is not easily corrected, precisely because of the evolving nature of social networks. When friends are updated about others' ratings and are influenced by them, online ratings may become path dependent. Managers should be aware of social networks' potential impacts on the value of their rating systems.

We conclude this paper by offering some caveats and limitations of our method. Valuable opportunities for future research are associated with these challenges. First, when users cannot perfectly observe the product's true quality even after consumption, as in the case of credence goods, it is possible for them to interpret their friends' ratings as quality signals. Although books (studied in this paper) and other information goods are generally considered to be experience goods in the literature (Shapiro and Varian 1999), it is difficult to completely rule out the possibility that social influence may arise from learning if there is postconsumption uncertainty in evaluating a book's quality. Our analysis of the moderating effects suggests that this concern is not significant in our context of online book ratings. Studies of other products should be careful about this. Valuable contributions can be made by future research to identify the exact mechanism through which social influence takes place.

Second, when users' evaluations are significantly different from their friends', they may choose not to post anything (Dellarocas 2006). Although this is also a type of social influence caused by social ties, its implication for rating systems could differ from the social nudge we find in this paper. Our data, however, do not allow us to investigate the significance of this type of influence directly.

Third, it is possible for pre- and postconsumption social influence to coexist in online product ratings. Our dyad-level analysis suggests that preconsumption social influence is less significant in book ratings. Future work could further differentiate and examine the relative importance of pre- and postconsumption social influence in online ratings.

Finally, there are some data-related limitations. (1) The data set was obtained from a Chinese social network, and how these results may be carried over to a different cultural setting requires some verification. We hope this study's methodological contribution will make such efforts easier. (2) We cannot examine the impact on sales. Moe and Trusov (2011) and Lee et al. (2014) make valuable contributions in this direction. (3) We were not able to present a full-fledged theoretical analysis about contingencies in social influence. Future research should extend our exploratory discussion about contingencies and establish a complete theoretical framework about online social influence.

Endnotes

[1] *Rating context* in this study refers to the virtual environment surrounding a user (reviewer) and the information therein.

[2] Parallel to this paper, Crandall et al. (2008) adopt a similar approach to examine similarity in the editing behavior of *Wikipedia* users.

[3] Focal-user ratings with no previous friends' ratings are excluded from the analysis.

[4] Rating situations with multiple friends' ratings are discussed in the online appendix (Figure A2).

[5] In the rest of this paper, "user" refers to registered users, as only registered users can have online friends and rate items.

[6] The mixed cases, as shown in panel C of Figure A2 in the online appendix, are analyzed separately in a robustness check.

[7] To alleviate the potential problem of nonnormality in some variables, we conduct our analysis with continuous variables log transformed. Our empirical results are robust and remain qualitatively the same with or without log transformation. Correlations are reported

in the online appendix. We also calculate the variance inflation factors (VIFs) according to Equation (1). VIFs of all variables are lower than 3, indicating that the independent variables do not suffer from serious multicollinearity issues (Kutner et al. 2004, Marquardt 1970).

[8] According to estimation results reported in column (2) of Table 2, the partial correlation between focal users' ratings and previous friend's ratings, controlling for other covariates, is 0.0259 before the formation of friend relationship. The number is 0.0753 after the friend relationship is formed, suggesting an increase of $0.0753/0.0259 - 1 = 191\%$. Meanwhile, if we look at the simple correlation between the two (from column (1)), the number is 0.2075 in the BEFORE cases and 0.2594 in the AFTER cases, an increase of 25%.

[9] Because of space limitations, details of the robustness check and additional analysis are provided in the online appendix.

[10] Both pre- and postconsumption influences are of social influence in online opinion reporting. We thank the anonymous reviewers for pointing out this distinction.

[11] We thank the anonymous reviewers for suggesting these tests.

[12] A detailed discussion about the inclusion of moderating variables is included in the online appendix (Table A9).

References

Angrist JD, Pischke JS (2010) The credibility revolution in empirical economics: How better research design is taking the con out of econometrics. *J. Econom. Perspect.* 24(2):3–30.

Antweiler W, Frank MZ (2004) Is all that talk just noise? The information content of Internet stock message boards. *J. Finance* 59(3):1259–1294.

Banerjee AV (1992) A simple model of herd behavior. *Quart. J. Econom.* 107(3):797–817.

Bapna R, Umyarov A (2015) Do your online friends make you pay? A randomized field experiment in an online music social network. *Management Sci.* 61(8):1902–1920.

Bikhchandani S, Hirshleifer D, Welch I (1992) A theory of fads, fashion, custom, and cultural change as informational cascades. *J. Political Econom.* 100(5):992–1026.

Bollen K, Pearl J (2013) Eight myths about causality and structural equation models. Morgan SL, ed. *Handbook of Causal Analysis for Social Research*, Handbooks Sociol. Soc. Res. (Springer, Dordrecht, Netherlands), 301–328.

Bolton GE, Katok E, Ockenfels A (2004) How effective are electronic reputation mechanisms? An experimental investigation. *Management Sci.* 50(11):1587–1602.

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

Brock WA, Durlauf SN (2001) Interactions-based models. Heckman JJ, Leamer E, eds. *Handbook of Econometrics*, Vol. 5 (North-Holland, Amsterdam), 3297–3380.

Brown JR, Ivković Z, Smith PA, Weisbenner S (2008) Neighbors matter: Causal community effects and stock market participation. *J. Finance* 63(3):1509–1531.

Burnkrant RE, Cousineau A (1975) Informational and normative social influence in buyer behavior. *J. Consumer Res.* 2(3):206–215.

Burtch G, Hong Y, Bapna R, Griskevicius V (2017) Stimulating online reviews by combining financial incentives and social norms. *Management Sci.*, ePub ahead of print March 3, https://doi.org/10.1287/mnsc.2016.2715.

Cai H, Chen Y, Fang H (2009) Observational learning: Evidence from a randomized natural field experiment. *Amer. Econom. Rev.* 99(3):864–882.

Campbell DT, Stanley JC (1963) *Experimental and Quasi-Experimental Designs for Research* (Houghton Mifflin, New York).

Chen Y, Xie J (2005) Third-party product review and firm marketing strategy. *Marketing Sci.* 24(2):218–240.

Chevalier JA, Mayzlin D (2006) The effect of word of mouth on sales: Online book reviews. *J. Marketing Res.* 43(3):345–354.

Chintagunta PK, Gopinath S, Venkataraman S (2010) The effects of online user reviews on movie box office performance: Accounting for sequential rollout and aggregation across local markets. *Marketing Sci.* 29(5):944–957.

Cialdini RB, Goldstein NJ (2004) Social influence: Compliance and conformity. *Annual Rev. Psych.* 55(1):591–621.

Cohen JB, Golden E (1972) Informational social influence and product evaluation. *J. Appl. Psych.* 56(1):54–59.

Conley TG, Udry CR (2010) Learning about a new technology: Pineapple in Ghana. *Amer. Econom. Rev.* 100(1):35–69.

Crandall D, Cosley D, Huttenlocher D, Kleinberg J, Suri S (2008) Feedback effects between similarity and social influence in online communities. *Proc.* 14*th ACM SIGKDD Internat. Conf. Knowledge Discovery Data Mining* (ACM, New York), 160–168.

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

Dellarocas C, Wood CA (2008) The sound of silence in online feedback: Estimating trading risks in the presence of reporting bias. *Management Sci.* 54(3):460–476.

Dellarocas C, Zhang XM, Awad NF (2007) Exploring the value of online product reviews in forecasting sales: The case of motion pictures. *J. Interactive Marketing* 21(4):23–45.

Deutsch M, Gerard HB (1955) A study of normative and informational social influences upon individual judgment. *J. Abnormal Soc. Psych.* 51(3):629–636.

Duan W, Gu B, Whinston AB (2008a) The dynamics of online word-of-mouth and product sales—An empirical investigation of the movie industry. *J. Retailing* 84(2):233–242.

Duan W, Gu B, Whinston AB (2008b) Do online reviews matter? An empirical investigation of panel data. *Decision Support Systems* 45(4):1007–1016.

Duflo E, Saez E (2003) The role of information and social interactions in retirement plan decisions: Evidence from a randomized experiment. *Quart. J. Econom.* 118(3):815–842.

Duggan M, Smith A (2013) Social media update 2013. Accessed December 15, 2014, http://www.pewinternet.org/2013/12/30/social-media-update-2013/.

Dulleck U, Kerschbamer R (2006) On doctors, mechanics, and computer specialists: The economics of credence goods. *J. Econom. Literature* 44(1):5–42.

Forman C, Ghose A, Wiesenfeld B (2008) Examining the relationship between reviews and sales: The role of reviewer identity disclosure in electronic markets. *Inform. Systems Res.* 19(3):291–313.

Foster G (2006) It's not your peers, and it's not your friends: Some progress toward understanding the educational peer effect mechanism. *J. Public Econom.* 90(8–9):1455–1475.

Godes D (2011) Commentary—Invited comment on "opinion leadership and social contagion in new product diffusion." *Marketing Sci.* 30(2):224–229.

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

Godes D, Silva JC (2011) Sequential and temporal dynamics of online opinion. *Marketing Sci.* 31(3):448–473.

Goes PB, Lin M, Au Yeung C (2014) "Popularity effect" in user-generated content: Evidence from online product reviews. *Inform. Systems Res.* 25(2):222–238.

Gu B, Park J, Konana P (2012) The impact of external word-of-mouth sources on retailer sales of high-involvement products. *Inform. Systems Res.* 23(1):182–196.

Hu N, Pavlou PA, Zhang J (2009) Why do online product reviews have a J-shaped distribution? Overcoming biases in online word-of-mouth communication. *Comm. ACM* 52(10):144–147.

Hu N, Bose I, Gao Y, Liu L (2011) Manipulation in digital word-of-mouth: A reality check for book reviews. *Decision Support Systems* 50(3):627–635.

Huang N, Hong Y, Burtch G (2017) Social network integration and user content generation: Evidence from natural experiments. *MIS Quart.* 41(4):1035–1058.

Irwin JR, McClelland GH (2001) Misleading heuristics and moderated multiple regression models. *J. Marketing Res.* 38(1):100–109.

Kuksov D, Xie Y (2010) Pricing, frills, and customer ratings. *Marketing Sci.* 29(5):925–943.

Kutner M, Nachtsheim C, Neter J (2004) *Applied Linear Regression Models*, 4th ed. (McGraw-Hill/Irwin, Boston).

Lazarsfeld PF, Merton RK (1954) Friendship as a social process: A substantive and methodological analysis. Berger M, Abel T, Page C, eds. *Freedom and Control in Modern Society* (Van Nostrand, New York), 18–66.

Lee YJ, Tan Y, Hosanagar K (2014) Do I follow my friends or the crowds? Examining informational cascades in online movie reviews. *Management Sci.* 61(9):2241–2258.

Lewis K (2011) The co-evolution of social network ties and online privacy behavior. Trepte S, Reinecke L, eds. *Privacy Online* (Springer, Berlin Heidelberg), 91–109.

Li X, Hitt LM (2008) Self-selection and information role of online product reviews. *Inform. Systems Res.* 19(4):456–474.

Liu D, Brass D, Lu Y, Chen D (2015) Friendships in online peer-to-peer lending: Pipes, prisms, and relational herding. *MIS Quart.* 39(3):729–742.

Liu Y (2006) Word of mouth for movies: Its dynamics and impact on box office revenue. *J. Marketing* 70(3):74–89.

Lu X, Ba S, Huang L, Feng Y (2013) Promotional marketing or word-of-mouth? Evidence from online restaurant reviews. *Inform. Systems Res.* 24(3):596–612.

Malthouse EC, Haenlein M, Skiera B, Wege E, Zhang XM (2013) Managing customer relationships in the social media era: Introducing the social CRM house. *J. Interactive Marketing* 27(4):270–280.

Manski CF (1993) Identification of endogenous social effects: The reflection problem. *Rev. Econom. Stud.* 60(3):531–542.

Marquardt DW (1970) Generalized inverses, ridge regression, biased linear estimation, and nonlinear estimation. *Technometrics* 12(3):591–612.

Marsden PV, Friedkin NE (1993) Network studies of social influence. *Sociol. Methods Res.* 22(1):127–151.

Mas A, Moretti E (2009) Peers at work. *Amer. Econom. Rev.* 99(1):112–145.

Mayzlin D, Dover Y, Chevalier J (2014) Promotional reviews: An empirical investigation of online review manipulation. *Amer. Econom. Rev.* 104(8):2421–2455.

McPherson M, Smith-Lovin L, Cook JM (2001) Birds of a feather: Homophily in social networks. *Annual Rev. Sociol.* 27:415–444.

Moe WW, Schweidel DA (2012) Online product opinions: Incidence, evaluation and evolution. *Marketing Sci.* 31(3):372–386.

Moe WW, Trusov M (2011) The value of social dynamics in online product ratings forums. *J. Marketing Res.* 48(3):444–456.

Moe WW, Schweidel DA, Trusov M (2011) Cutting through online chatter: White noise or resonating insights? *Sloan Management Rev.* 53(1):14–16.

Moretti E (2011) Social learning and peer effects in consumption: Evidence from movie sales. *Rev. Econom. Stud.* 78(1):356–393.

Nielsen (2014) The U.S. digital consumer report. Accessed December 15, 2014, http://www.nielsen.com/us/en/insights/reports/2014/the-us-digital-consumer-report.html.

Pavlou PA, Gefen D (2004) Building effective online marketplaces with institution-based trust. *Inform. Systems Res.* 15(1):37–59.

Resnick P, Zeckhauser R (2002) Trust among strangers in Internet transactions: Empirical analysis of eBay's reputation system. Baye MR, ed. *The Economics of the Internet and e-Commerce*, Adv. Appl. Microeconomics, Vol. 11 (JAI Press, Amsterdam), 127–157.

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

Salganik MJ, Dodds PS, Watts DJ (2006) Experimental study of inequality and unpredictability in an artificial cultural market. *Science* 311(5762):854–856.

Schlosser AE (2005) Posting versus lurking: Communicating in a multiple audience context. *J. Consumer Res.* 32(2):260–265.

Shapiro C, Varian HR (1999) *Information Rules: A Strategic Guide to the Network Economy* (Harvard Business Review Press, Boston).

Snijders TAB, Steglich CEG, Schweinberger M (2006) Modeling the coevolution of networks and behavior. van Montfort K, Oud J, Satorra A, eds. *Longitudinal Models in the Behavioral and Related Sciences* (Routledge, New York), 41–72.

Soetevent AR (2006) Empirics of the identification of social interactions; An evaluation of the approaches and their results. *J. Econom. Surveys* 20(2):193–228.

Steglich C, Snijders TAB, Pearson M (2010) Dynamic networks and behavior: Separating selection from influence. *Sociol. Methodology* 40(1):329–393.

Summers LH (1991) The scientific illusion in empirical macroeconomics. *Scandinavian J. Econom.* 93(2):129–148.

Susarla A, Oh JH, Tan Y (2012) Social networks and the diffusion of user-generated content: Evidence from YouTube. *Inform. Systems Res.* 23(1):23–41.

Thaler RH, Sunstein CR (2008) *Nudge: Improving Decisions About Health, Wealth, and Happiness* (Yale University Press, New Haven, CT).

Trusov M, Bodapati AV, Bucklin RE (2010) Determining influential users in Internet social networks. *J. Marketing Res.* 47(4):643–658.

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

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

Watts DJ, Dodds PS (2007) Influentials, networks, and public opinion formation. *J. Consumer Res.* 34(4):441–458.

Wu F, Huberman BA (2008) How public opinion forms. Papadimitriou C, Zhang S, eds. *Internet and Network Economics*, Lecture Notes Comput. Sci., Vol. 5385 (Springer, Berlin Heidelberg), 334–341.

Xu SX, Zhang XM (2013) Impact of Wikipedia on market information environment: Evidence on management disclosure and investor reaction. *MIS Quart.* 37(4):1043–1068.

Yin D, Mitra S, Zhang P (2015) When do consumers value positive versus negative reviews? An empirical investigation of confirmation bias in online word of mouth. *Inform. Systems Res.* 27(1):131–144.

Yoganarasimhan H (2012) Impact of social network structure on content propagation: A study using YouTube data. *Quant. Marketing Econom.* 10(1):111–150.

Zhang J (2010) The sound of silence: Observational learning in the U.S. kidney market. *Marketing Sci.* 29(2):315–335.

Zhang J, Liu P (2012) Rational herding in microloan markets. *Management Sci.* 58(5):892–912.

Zhang XM, Wang C (2012) Network positions and contributions to online public goods. *J. Management Inform. Systems.* 29(2):11–40.

Zhang XM, Zhu F (2011) Group size and incentives to contribute: A natural experiment at Chinese Wikipedia. *Amer. Econom. Rev.* 101(4):1601–1615.

Zhu F, Zhang XM (2010) Impact of online consumer reviews on sales: The moderating role of product and consumer characteristics. *J. Marketing* 74(2):133–148.