Management Science

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

Monetary Incentives and Knowledge Spillover: Evidence from a Natural Experiment

Jing Wang , Gen Li , Kai-Lung Hui

informs
http://pubsonline.informs.org/journal/mnsc

MANAGEMENT SCIENCE
Articles in Advance, pp. 1–24
ISSN 0025-1909 (print), ISSN 1526-5501 (online)

Monetary Incentives and Knowledge Spillover: Evidence from a Natural Experiment

Jing Wang,[a] Gen Li,[b,*] Kai-Lung Hui[a]

[a] Department of Information Systems, Business Statistics, and Operations Management, School of Business and Management, Hong Kong University of Science and Technology, Hong Kong, China; [b] Department of Information Management and Information Systems, School of Management, Fudan University, Shanghai 200433, China
*Corresponding author

Contact: jwang@ust.hk, https://orcid.org/0000-0002-8877-7134 (JW); gli@fudan.edu.cn, https://orcid.org/0000-0001-9493-3976 (GL); klhui@ust.hk, https://orcid.org/0000-0002-7074-1176 (K-LH)

Received: August 10, 2019
Revised: May 3, 2020; December 5, 2020
Accepted: February 26, 2021
Published Online in Articles in Advance: July 1, 2021

https://doi.org/10.1287/mnsc.2021.4048

Abstract. We examine how the introduction of monetary incentives by a knowledge-sharing platform affects the nonrewarded knowledge activity on the platform. Our setting is a question-and-answer platform that provides monetary incentives for holding live talks. Using a combination of coarsened exact matching and difference-in-differences estimation techniques, we find that the launch of the paid feature creates a positive spillover effect on the hosts' free contributions, specifically, 9.4%–40.8% more answers in the short run when compared with nonhosts. The paid feature did not result in any significant change in the quality of answers. We suggest reputation building is one plausible mechanism underlying the spillover. Additional analyses reveal the spillover effect is negative for short-lived hosts in the long run, indicating possible crowding out of free contributions. The positive spillover effect for long-lived hosts lasts longer but is also reduced over time. Our findings suggest that introducing monetary incentives can be a viable business model for knowledge platforms to stimulate user contributions in both the paid and related unpaid activities. Yet platform owners should be cautious about the potential negative spillover after users stop contributing to the paid activity and develop effective strategies to maintain users' long-term interests in the paid program.

History: Accepted by Kartik Hosanagar, information systems.

Supplemental Material: The online appendix and data files are available at https://doi.org/10.1287/mnsc.2021.4048.

Keywords: online Q&A communities • monetary incentives • spillover effect • reputation building • crowding out

1. Introduction

Knowledge platforms such as content-sharing sites (e.g., WordPress, YouTube), online encyclopedia (e.g., Wikipedia), and question-and-answer (Q&A) communities (e.g., Quora) are becoming popular on the internet. The sustainability of these platforms relies on members' participation and, more importantly, contribution of knowledge. To encourage more and better contributions, some platforms have started to introduce monetary incentives to contributors. For example, YouTube shares about half of the advertising revenue generated from a video with its creator.[1] WordPress shares the majority of advertising revenue with its bloggers on a per-impression basis.[2]

Another, more direct way of paying knowledge contributors is to make users pay for consuming their contents. For example, the massive open online course (MOOC) platform Udemy allows experts to create courses and charge a tuition fee from students.[3] One of the most well-known Q&A websites in China, Zhihu, offers the Zhihu Live service, which allows contributors to hold live talks and collect entrance fees from participating listeners.[4] The popularity of these platforms and services indicates that monetary incentives can help facilitate knowledge contribution in the rewarded activities (offering online courses and holding live talks). An important question, however, is whether such monetary incentives affect users' contribution to nonpaid ("free") knowledge activities.

Answering this question is important because consumers benefit from both paid and nonpaid knowledge. If paying for MOOC causes instructors to stop sharing free lecture videos or conducting free seminars or if rewarding live talks causes contributors to stop answering questions or authoring free articles on Zhihu, then facilitating monetary rewards may not increase the overall contribution of knowledge to the community. If, however, the payment causes contributors to provide more free knowledge activities, then its benefit to knowledge contribution is unambiguously positive.

Several distinct mechanisms suggest that monetary incentives could have a negative spillover effect on users' contribution to free knowledge activities. First, because contributors have finite time and effort, they need to allocate their limited resources across different knowledge activities (Holmstrom and Milgrom 1991). If they spend time and effort on rewarded activities, they may have less capacity for free activities, leading to an effort distortion effect; that is, contributors spend more effort on rewarded activities and less effort on nonrewarded activities (Brüggen and Moers 2007, Kachelmeier et al. 2008). Second, monetary rewards may create a crowding-out effect (Titmuss 1970, Deci et al. 1999). Knowledge contribution may be driven by motivations such as helping others, knowledge self-efficacy, or reciprocal benefits. Offering monetary compensation could undermine such motivations, making free contributions less appealing to contributors. This could lead to a net reduction in free knowledge contribution.

On the other hand, monetary incentives could have a positive spillover effect. First, consumers must pay before accessing the compensated knowledge, meaning contributors need to build a reputation before they can attract paying audience. By making free and publicly accessible contributions, contributors can send a credible signal about their capability and raise awareness of themselves in the community. Second, contributors could use free contributions as an avenue to explicitly promote and refer readers to their paid knowledge. Third, the monetary compensations received by contributors may induce them to reciprocate (Gouldner 1960, Falk and Fischbacher 2006). Contributors may feel obliged to behave more nicely and return the "favor" by making more free contributions to the community.

Given that monetary incentive can variously create negative and positive spillover effects, its net impact on users' contribution to nonpaid activities is an important empirical question.[5] In this paper, we study the experience of Zhihu, a leading player in the "pay-for-knowledge" industry in China. In May 2016, Zhihu launched Zhihu Live, a service that allows users to monetize their knowledge via conducting interactive live talks. We empirically investigate the spillover effect of this new service on the hosts' contributions to unpaid activity of answering questions. We use the difference-in-differences (DID) strategy to compare the before–after changes in free contributions by live hosts (users who hosted at least one live talk) vis-à-vis nonhosts (users who have not hosted any live talk). To address potential self-selection bias, we use the coarsened exact matching (CEM) method to construct matched sets of live hosts and nonhosts with comparable user characteristics.

Our DID estimation results using the CEM sample show that monetary incentive causes a positive spillover effect on users' free contribution to answering questions. Compared with nonhosts, live hosts contributed 9.4%–40.8% more answers in the six months after the cutoff dates than in the six months before.[6] The paid feature did not result in any significant change in the quality of answers regardless of whether quality is measured by the number of characters used or the number of vote-ups received.

To validate our matching and estimation procedures, we perform a falsification test by setting the "placebo" cutoff dates to six months before the true cutoff dates. The absence of positive spillover effect in this exercise provides indicative evidence that our procedures are appropriate. We also employ a lookahead matching strategy that matches existing hosts to future hosts and find a similar positive spillover effect. The consistent results suggest that our findings are not driven by selection on unobserved user characteristics.

We find that reputation building is one plausible explanation of the positive spillover.[7] First, we show that a host's higher reputation can translate into more revenue from the live talks via both higher entrance fees and more listeners. Second, we demonstrate that the positive spillover effect is stronger among live hosts with lower prior reputation, who should have greater needs to signal their competence and raise public awareness. We provide direct evidence that our finding persists in view of competing explanations, such as explicit promotion and reciprocity.

We conduct additional analyses to explore the contributions of live hosts in the long run. Overall, we do not observe any long-term spillover effect of the paid feature on either the quantity or the quality of contributed answers. This long-run result holds for both hosts with low prior reputation and those with high prior reputation. However, if we only consider short-lived hosts, defined as users who held talks for a short period and stopped afterward, the long-term spillover effect is actually negative, indicating that the provision of monetary incentive may crowd out nonmonetary motivations of contributors (Deci 1971). We rule out two alternative explanations that may drive the negative long-run spillover effect for short-lived hosts: user exit and expectation disconfirmation. The positive spillover effect for long-lived hosts, who continued holding talks for a longer period, lasted longer, although it also reduced over time.

This paper makes three important contributions. First, it extends our knowledge of how monetary incentives affect online contributions. Although many studies have explored the direct effect of monetary incentives on the rewarded contributions (product reviews, blog posts, answering questions, etc.), little research has studied its spillover effects on the quantity and quality of related nonrewarded contributions. Our finding, that offering monetary incentives to a paid activity motivates users' contributions to a major unpaid activity, is encouraging in that the overall contribution of knowledge to the community is enhanced. Furthermore, this increase in contribution quantity does not come at the expense of reduction in average quality. The novel spillover benefit uncovered in our research supports the use of monetary incentives to stimulate user contribution in online communities.

Second, we advance reputation building as a plausible underlying mechanism causing the spillover. In online Q&A communities, reputation management is an important motivation for users' voluntary contributions (e.g., Wasko and Faraj 2005, Yu et al. 2007, Lou et al. 2013). In the workplace, enhancing reputation through knowledge sharing can be beneficial when contributors are recognized as having expertise and earn respect and appreciation, which might turn into organizational rewards, such as better evaluation, increased pay, or job promotion. Recent research has found evidence for a positive linkage between users' online contribution and offline career advancement in the open source software (OSS) community (Hann et al. 2013) and programming Q&A platforms (Xu et al. 2020). We contribute to this literature by showing that monetizing contributions alone (cf. off-line rewards or career advancements) can motivate users to build their reputations on online knowledge-sharing platforms.

Third, this study extends the limited literature on the long-term effect of monetary intervention (e.g., Meier 2007, Mochon et al. 2016) by investigating its heterogeneous spillover effects among different types of contributors. Our results show that the provision of monetary incentives may crowd out the intrinsic and image motivations, leading to a potential negative spillover effect for contributors who participate in the paid activity for a short period and stop afterward. However, for contributors who engage in the paid activity for a longer period, the positive spillover effect lasts longer although it is considerably weaker over time. The nuanced understanding of long-term effects offers valuable design implications on the use of monetary incentives.

The remainder of this paper is organized as follows. Section 2 reviews related literature. Section 3 describes our research setting and data. Section 4 details the analysis design and identification strategy. Section 5 presents the empirical model and results. Section 6 explores the underlying mechanism. Section 7 reports long-term effects. Section 8 discusses and concludes the paper.

2. Literature Review

Why people voluntarily contribute their knowledge in online communities is an important topic that attracts extensive research in the literature. The factors leading to such contribution may include enjoyment in helping others, knowledge self-efficacy, reciprocal benefits, reputation building, and career improvement (Wasko and Faraj 2005, Lin 2007, Yu et al. 2007, Lou et al. 2013, Chen et al. 2018). Several empirical studies show that nonmonetary rewards encourage contribution (Li et al. 2012, Khansa et al. 2015, Goes et al. 2016). Recent research, however, has started to examine the effects of tangible rewards, such as monetary compensation or career opportunities, on the quantity and quality of user contribution (Chen et al. 2010, Hsieh et al. 2010, Xu et al. 2020). Our research extends this line of work by studying how monetary incentives enhance users' motivation to build their reputation through nonpaid contributions. We review the broad literature on monetary incentives and reputation building.

2.1. Monetary Incentives and User Behavior

Our study builds on the vast literature of how individuals respond to extrinsic rewards. In general, agency theory suggests that financial compensations can incentivize more effort and better performance. However, the presence of monetary rewards may crowd out individuals' intrinsic and image motivations (Titmuss 1970, Frey and Oberholzer-Gee 1997, Deci et al. 1999, Bénabou and Tirole 2006, Ariely et al. 2009).[8] The net effect on user behavior when incentives are in place may depend on the magnitude of monetary rewards. Gneezy and Rustichini (2000) show that subjects who were offered a small compensation invested less effort than those who were not compensated. When compensation was offered, a larger amount resulted in more effort. Nevertheless, after removing the incentives, user behavior could fall below the preintervention period. For example, Meier (2007) shows that a matching incentive increases donation in the short run but decreases donation in the long run. The long-term effect can be positive, however, if incentives can foster the development of good habits (Charness and Gneezy 2009, Mochon et al. 2016).

Recently, there has been a growing body of studies on the role of monetary incentives on users' online contributions. Monetary rewards can increase (e.g., Burtch et al. 2017) or decrease (e.g., Sun et al. 2017) the volume of online reviews and improve (e.g., Stephen et al. 2012), undermine (e.g., Khern-am-nuai et al. 2018), or have no effect (e.g., Wang et al. 2012, Burtch et al. 2017) on review quality. Sun and Zhu (2013) find that sharing ad revenue increases the number and quality of blog posts and induces contributors to shift their content toward more popular topics. Fershtman and Gandal (2007) find that more commercially oriented OSS projects induce greater effort investment from contributors. Specific to Q&A communities, the provision of monetary rewards increases the quantity of answers (Hsieh et al. 2010), but the empirical evidence on answer quality is mixed. Chen et al. (2010) and Hsieh et al. (2010) show that paying more elicits longer but not better answers, whereas Harper et al. (2008) find that paying more leads to answers of higher quality. In an online health community, Zhang et al. (2017) find that the introductory incentives increase new physicians' contribution quantity and quality during the policy window, but the effect is negative after the policy window.

Instead of studying how monetary incentive directly affects the contribution to the rewarded activities, we focus on its spillover effects on the quantity and quality of related nonrewarded activities. The closest work to ours is Kuang et al. (2019), which studies the spillover effects of financial incentives on the quantity of nonincentivized user engagements. They did not find any spillover effect on the hosts' free answer contributions. Here, we examine both the quantity and quality of free answer contributions, use a more extensive data set and various identification strategies, and explore the underlying mechanisms causing the spillover effect. We also investigate the long-term effects of the monetary incentives. To our knowledge, this is the first study to investigate such spillover effects in both the short and long term.

2.2. Reputation Building, Signaling Quality, and Raising Awareness

Our research is also related to the literature on reputation building. At the corporate level, the concept of reputation encompasses both the public awareness and visibility and the overall assessment or judgment of the company (Barnett et al. 2006, Lange et al. 2011). Companies with a superior reputation can achieve competitive advantages, such as enjoying higher demand, charging premium prices, and attracting potential investors. Firms may offer free samples, advertise, or generate media publicity to build their reputation. It is conceivable that, in a knowledge community, making free contributions can help enhance user reputation by signaling quality and raising awareness.

In an organizational setting, people use a variety of tactics to signal ability and gain recognition in the hope that they can bring tangible benefits, such as better performance appraisal (e.g., Wayne and Ferris 1990), pay raises (e.g., Weiss 1995), and career success (e.g., Judge and Bretz 1994). Signaling quality is a primary incentive for developers' voluntary contributions in OSS and online programming Q&A communities (Lerner and Tirole 2002, Hann et al. 2013, Xu et al. 2020). Participating in OSS projects or answering programming questions sends a credible signal of one's skills and capability, which may bring tangible benefits, such as increased pay or job opportunities. Besides this, users in online communities may also make free contributions to seek attention and raise awareness. For example, Rui and Whinston (2012) and Tang et al. (2012) suggest that the desire for attention and exposure is a major incentive for content contribution in social media.

Our paper extends this literature by showing how monetary incentives affect individuals' motivation to build reputation within the same online community. We also scrutinize the delicate choices of users in variously making quantity and quality contributions to free knowledge in both the short and long terms. Such free contribution is instrumental to reputation building in an online setting in which users do not physically interact with each other.

3. Research Setting and Data

3.1. Study Context

Our empirical setting is Zhihu, a popular online Q&A community in China. Zhihu was launched in January 2011 and quickly became one of the most frequently visited websites by Chinese users. Users on Zhihu can ask and answer questions, write articles, and comment and vote on the answers and articles. As of December 2020, Zhihu had more than 370 million registered users and had accumulated 44 million questions and 260 million answers.[9] Since the beginning, all contributions on Zhihu can be freely accessed by its members. In May 2016, the platform launched Zhihu Live, a service that enables users to conduct live talks and collect entrance fees from listeners attending the talks. It was Zhihu's first monetization attempt to compensate contributors for their knowledge sharing.

With Zhihu Live, any user can offer a live talk of one to two hours within Zhihu's mobile app.[10] The speaker can post the talk's details, including time, price, title, and brief introduction, before the live session. Figure 1 shows an example of a live talk description. Other users can pay a small entrance fee (9.99 RMB, 19.99 RMB, etc.) set by the speaker to join the talk. The live talks are usually delivered via discontinuous short audio messages, photos, text, or slides. Throughout the live session, listeners can ask questions and interact with the speaker in real time. Figure 2 presents a snapshot of an ongoing live talk. After the talk ends, interested users can still attend the talk but cannot interact

Figure 1. (Color online) Sample Zhihu Live Talk Description

with the host. Zhihu does not impose any constraint on the number of listeners that can join a single talk.

As of May 2018, Zhihu Live hosted about 6,000 live talks, generating 150 million RMB revenue from more than seven million paying listeners. On average, each speaker delivered sessions that were 65 minutes long and received an hourly reward of around 11,000 RMB. Zhihu does not make any official announcement of newly launched live talks. Instead, it recommends the talks of the speakers to potential interested users in the "recommendation" column based on their past interactions, for example, users who have upvoted the answers or articles contributed by the speakers before. In addition, Zhihu pushes users' activities to their followers directly in the "follow" column. If a user follows a live talk speaker, then the speaker's recent activities, including holding live talks, are pushed to the user. Users can also see the detailed profiles of live talk speakers on Zhihu, including all of their past activities (e.g., posting questions, answers, and articles) on the platform, the cumulative number of vote-ups received, and detailed information of all the talks held or to be held by them. Therefore, making free contributions is an effective means for live talk speakers to raise their visibility and cultivate more potential listeners.

3.2. Data

Our data period ends in May 2018, two years after the launch of Zhihu Live in May 2016. The sample consists of all 2,240 live hosts who held at least one live talk before the end of data collection and a total of 4,467 nonhosts using a snowball sampling technique. Specifically, we started with one random live host and collected all that host's followees. We then collected all followees of these first-degree followees. We repeated these steps several times until we reached 30,000 users. Finally, we randomly selected 5,000 users from this list, out of which 4,467 are nonhosts. Table H1 in the online appendix presents a comparison between the live hosts and nonhosts in terms of activity level, which confirms that this sampling procedure allows us to construct a group of nonhosts who are comparable to the live hosts.[11] Both groups, however, are much more active than the average

Figure 2. (Color online) Snapshot of an Ongoing Zhihu Live Talk

population of Zhihu. This is expected as user activities in online Q&A communities often follow a power-law distribution in which a small number of users contribute the majority of the content.

For each user, we recorded the user type (personal or organization account) and all user activity data on Zhihu, including hosting live talks, asking and answering questions, and writing articles.[12] For each activity, we recorded the user ID, time of the activity, title, content, number of vote-ups, and the associated tags. To eliminate potential confounds from the marketing activities of organization accounts, we remove all 183 organization accounts from our analysis. We then constructed a panel data set with user as the cross-sectional unit and month as the time unit.

The dependent variables are the quantity and quality of users' free knowledge contribution in terms of answering questions.[13] We measure contribution quantity by the number of answers provided by user i in month t. We measure contribution quality by the average number of characters per answer and the average number of vote-ups per answer for the answers contributed by user i in month t.[14] The number of characters reflects the amount of effort that the users put into answering the questions. The number of vote-ups captures readers' assessment of the quality of the answers.

Table 1 presents the variables' definitions and descriptive statistics. The distributions of the dependent variables are highly skewed with a small number of users having extreme values. In the main analysis, we log transform these variables to reduce the excessive influence of outliers.[15] Accordingly, the estimated effects can be interpreted as the elasticity. We use a Poisson regression model to corroborate the estimations in a robustness check.

4. Design and Identification Strategy

In our data set, the time when users began to hold live talks is spread over the two years after the introduction of Zhihu Live. Given that some live hosts could have been aware of the policy much later than the date it was introduced, using a single cutoff date would bias the estimate downward. Therefore, we adopt a staged design to examine the effect of the monetary incentive. Specifically, we partition the 2,240 live hosts into four groups. Group G1 consists of 469 live hosts who held their first talks within the first six months after Zhihu Live was introduced. Group G2 consists of 670 live hosts who held their first talks in the second six-month period. Group G3 consists of 698 live hosts who held their first talks in the third six-month period. Group G4 consists of 403 live hosts who held their first talks in the fourth six-month period. For each group, the "before" period is the six-month window before the cutoff date of the group. The "after" period is the six-month window after the cutoff date. This staged design allows us to accommodate the heterogeneous exposure of live hosts to the

Table 1. Descriptive Statistics

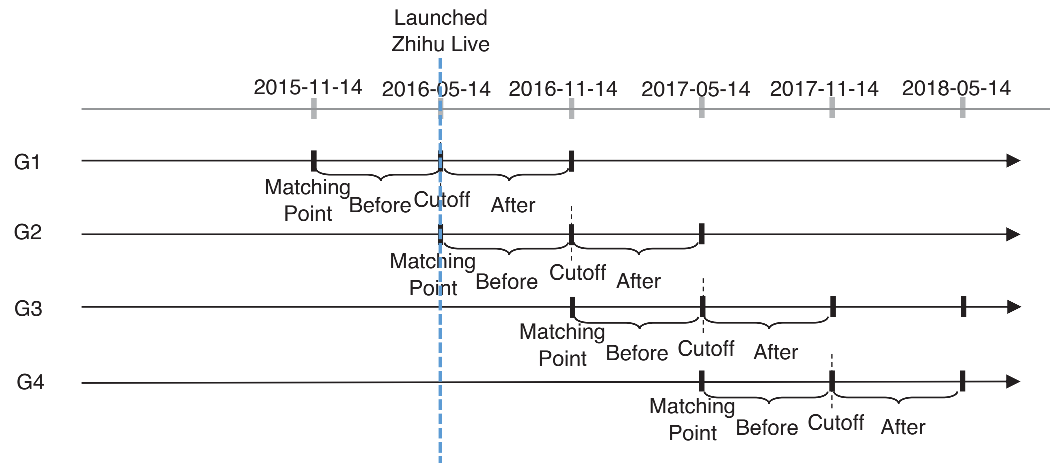
Variable	Description	Mean	Standard deviation	Minimum	Maximum
$Treatment_i$	Whether user i is in the treatment group of G1–G4	0.333	0.471	0	1
$Tenure_i$	Number of months since user i's first answer contribution (by the end of the data collection)	35.149	21.236	1	89
$Num\_Answers_{it}$ (number of answers)	The number of answers contributed by user i in month t	4.985	17.327	0	1,068
Num_Chars_{it} (number of characters)	The average number of characters per answer contributed by user i in month t	551.578	940.477	0	58,427
$Num\_Voteups_{it}$ (number of vote-ups)	The average number of vote-ups per answer contributed by user i in month t	323.262	1,621.104	0	170,143

Zhihu live talk feature.[16] Figure 3 summarizes the design of the four live host groups.

We use DID, a quasi-experimental design, to identify the effect of the paid feature on the hosts' free answer contributions. One potential concern of DID in our setting is that the "live host" treatment may not be randomized. There could be self-selection bias as the decision to host live talks is endogenous. To address this concern, we use CEM to construct comparable treatment and control groups (Blackwell et al. 2009). The principle of CEM is to "coarsen" each variable according to user-defined cut points or an automatic binning algorithm and perform exact matching on the coarsened data (known as stratum). It has several advantages over other commonly used matching methods, such as propensity score matching (PSM), as it does not make any assumption about the underlying data-generation process; greatly reduces imbalance, model dependence, and estimation errors; and is easy to understand and implement (Iacus et al. 2012). The use of CEM has gained increasing popularity in information systems research (e.g., Overby and Forman 2014). To fully exploit the data, we apply CEM weighted matching using all observations but with control users having different weights determined by the ratio of treatment and control units in each stratum.

For each treatment group, that is, G1–G4, we use all nonhosts to construct a matched sample. We set the matching points to be six months prior to the cutoff dates and perform CEM matching for each treatment group using data *up to* the corresponding matching

Figure 3. (Color online) Live Host Group Design



point. For example, the matched sample was constructed from the nonhosts' data up to November 14, 2015, for group G1; May 14, 2016, for group G2; and so on. We restrict the analysis to users who provided at least one answer prior to the matching point to remove users who were new or not predisposed to contributing answers on Zhihu. Finally, we perform CEM weighted matching on the live hosts and nonhosts using the log average number of answers per month, log average number of characters per answer, log average number of vote-ups per answer, and tenure up to the matching point.[17]

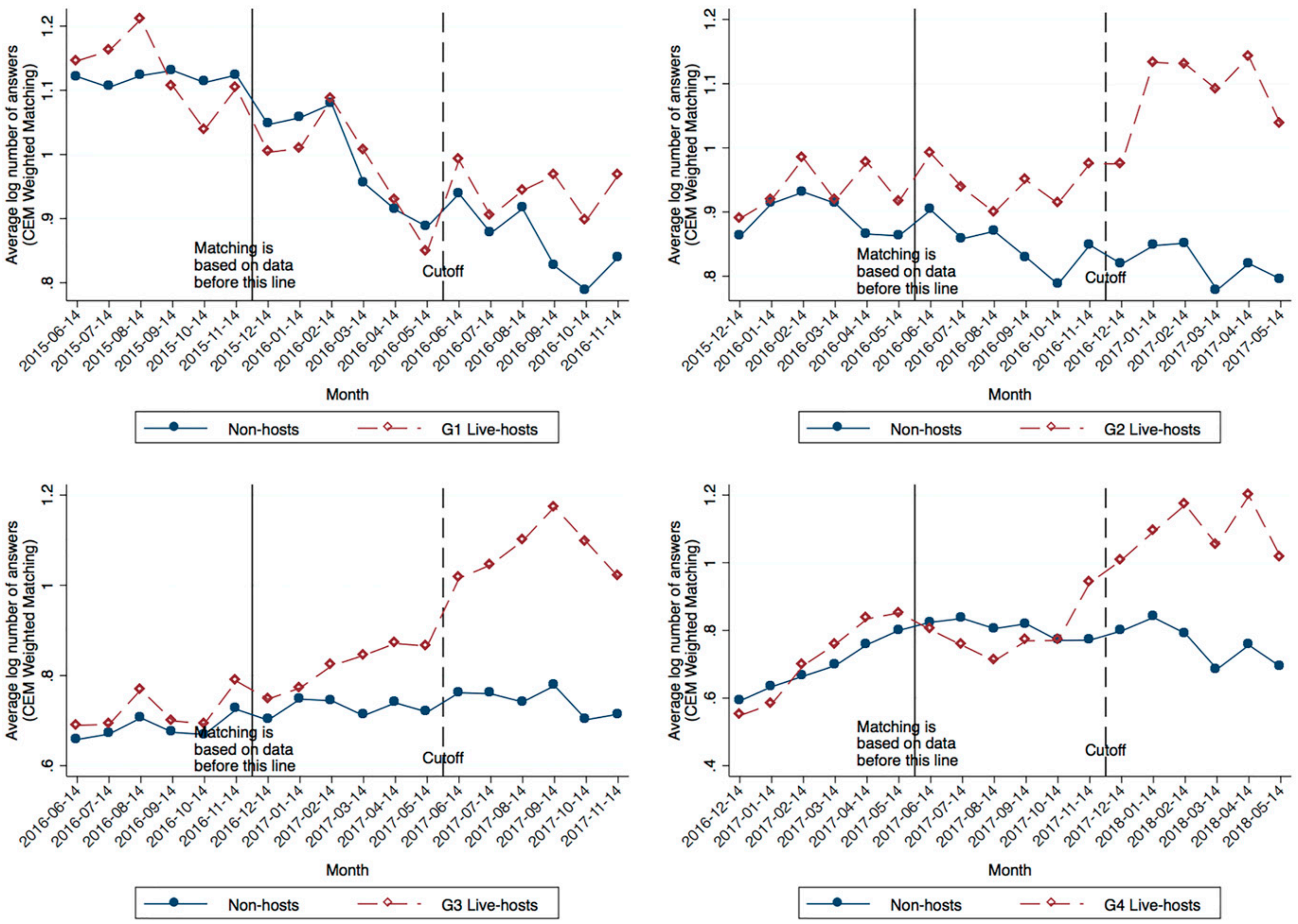
We set the matching points to be six months before the cutoff dates to accommodate the possibility that some live hosts could have strategically altered their behavior before hosting their first live talks. The earlier matching points also help us assess the possibility of overfitting, that the control group could have been constructed from users who exhibit similar values in the matching variables in the earlier periods but are fundamentally different. The period from the matching point to the cutoff date allows us to check the validity of the matching, specifically, whether the matched pairs still behave similarly after the matching period. We can use this novel matching strategy because we track all user activities on Zhihu throughout their entire lifetime, which often spans a number of years.

5. Empirical Analysis

5.1. Model-Free Evidence

To visualize the effect of the paid feature, we present model-free evidence using the average values of the log-transformed dependent variables for the matched live hosts and nonhosts. Figure 4 presents the results on answer quantity. The matched live hosts and nonhosts contributed similar numbers of answers before the matching point, which is expected as the matching was performed based on users' answer contributions up to this point. Their contributions followed similar trends afterward, which adds confidence to the validity of our matching approach. It is worth noting that the answer contributions of treatment groups G2–G4 started to increase a few months before the cutoff dates,

Figure 4. (Color online) Average Answer Quantity, Staged Design



Notes. The solid vertical line represents the matching point of each group. The dashed vertical line represents the cutoff date of each group.

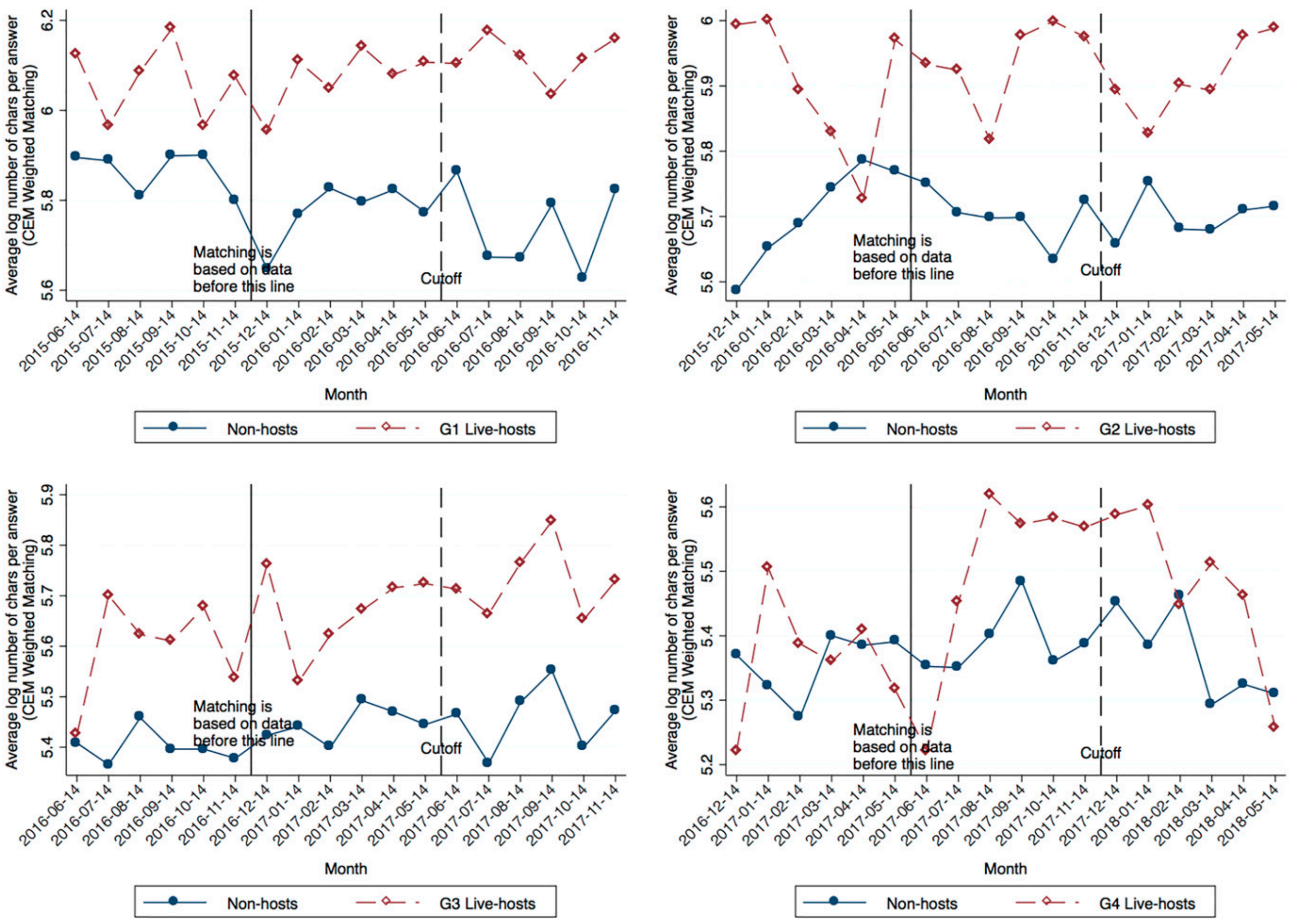
possibly because some live hosts wanted to build their reputation in advance for their live talks. However, such incentive and strategic behavior should not be observed for live hosts in G1 because users were not aware of the paid feature before the cutoff date for G1.[18] Later, in Section 5.4.3, we shift the matching points backward to one year instead of six months before the cutoff dates and find a similar pattern.

After the cutoff dates, nonhosts still followed the original trends in answer contributions, but the answer contributions of live hosts increased dramatically. This indicates that the paid feature motivates live hosts to contribute more to the unpaid activity of answering questions in the short term. Furthermore, such effect seems to be stronger among live hosts who held their first talks at a later time (i.e., treatment groups G2–G4). One potential concern with the DID strategy is that the behavior of nonhosts may change because of the introduction of Zhihu Live, which could compromise the integrity of the control group and the validity of the estimated effect. However, the model-free graphs in Figure 4 suggest that this is unlikely to be the case.

We next consider answer quality. Figures 5 and 6 present the quality trends for matched live hosts and nonhosts in terms of the average log number of characters and average log number of vote-ups per answer. Notably, the quality trends are more volatile than the quantity trends, possibly because most users contribute a small number of answers per month and the set of users for quality measurements varies over months as many users do not contribute answers every month. Overall, the CEM matching helps yield matched samples of live hosts and nonhosts who contribute answers of comparable quality. The paid feature does not produce any noticeable change in answer quality regardless of whether answer quality is measured by the number of characters used or the number of vote-ups received.

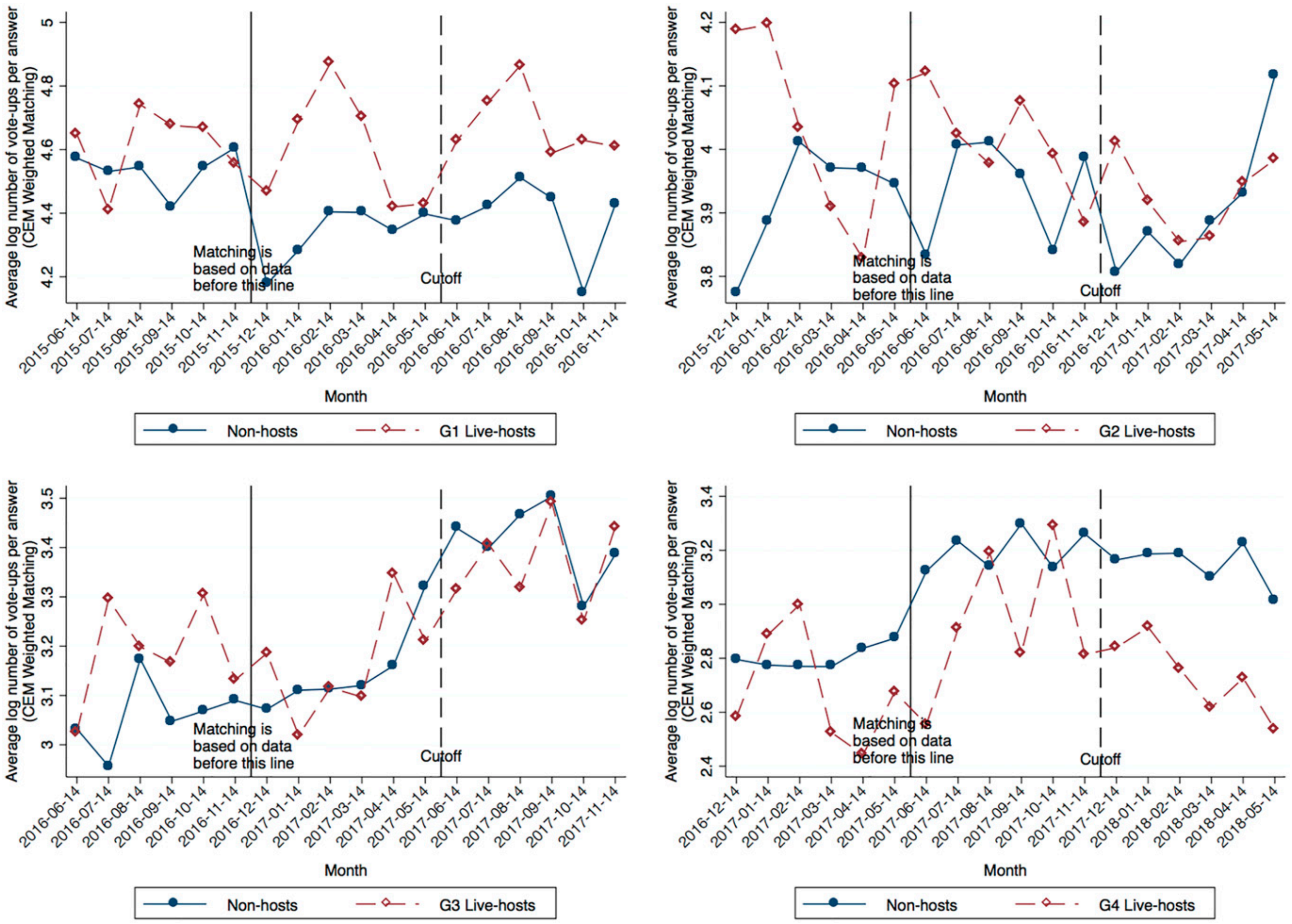
The model-free graphs offer preliminary evidence that the paid feature has a positive impact on the quantity of free answers provided by live hosts, but its

Figure 5. (Color online) Average Answer Quality (Number of Characters), Staged Design



Notes. The solid vertical line represents the matching point of each group. The dashed vertical line represents the cutoff date of each group.

Figure 6. (Color online) Average Answer Quality (Number of Vote-ups), Staged Design



Notes. The solid vertical line represents the matching point of each group. The dashed vertical line represents the cutoff date of each group.

effect on quality is not apparent. Next, we conduct formal statistical analyses to examine the spillover effects.

5.2. The Empirical Model

Our main model for the empirical analysis is

$$y_{it} = \beta_0 + \beta_1 Treatment_i \times After_t + u_i + \tau_t + \epsilon_{it}, \quad (1)$$

where y_{it} denotes the quantity or quality of user i's free answer contribution. The variable $Treatment_i$ is a dummy that equals one if user i is a live host (i.e., if user i is in treatment group G1–G4) and zero if user i is a nonhost (i.e., if user i is in the control group). For each group, the investigation period is the six months before and after the cutoff date as shown in Figure 3. $After_t$ equals zero before the cutoff date of each group and one otherwise. u_i captures time-invariant user fixed effects. τ_t captures time fixed effects. ϵ_{it} captures any idiosyncratic random errors. With this specification, the coefficient β_1 is our DID estimator. It captures the net impact of the paid feature on live hosts' contribution to free answers by controlling for user heterogeneity and time-specific shocks affecting the contribution.

5.3. Main Results

We report the main results obtained using the staged design in this section. Column (1) of Table 2 presents the regression results of the matched samples with the quantity of answer contribution as the dependent variable. The coefficient of $Treatment_i \times After_t$ is positive and statistically significant for live hosts in all four treatment groups. Compared with matched nonhosts, live hosts in G1 contributed 9.4% more answers in the six months after the cutoff date.[19] This effect increases to 18.6%, 27.0%, and 40.8% for G2, G3, and G4, respectively. These results indicate that the paid feature has motivated live hosts to contribute more to a related noncompensated activity, that is, answering questions, in the same online community. In addition, the later live hosts held their first talks, the more pronounced the positive spillover effect.[20] We explore the reason in Section 6.1.

Table 2. Main Regression

CEM weighted matching		Quantity (number of answers) (1)	Quality (number of characters) (2)	Quality (number of vote-ups) (3)
G1	Treatment × after	0.090**	0.112**	−0.004
		(0.040)	(0.055)	(0.068)
	Number of observations	22,380	11,985	11,985
	Number of users	1,865	1,646	1,646
	Adjusted R^2	0.704	0.480	0.628
G2	Treatment × after	0.171***	0.032	0.002
		(0.033)	(0.054)	(0.060)
	Number of observations	29,916	14,652	14,652
	Number of users	2,493	2,160	2,160
	Adjusted R^2	0.709	0.491	0.666
G3	Treatment × after	0.239***	0.057	0.068
		(0.041)	(0.056)	(0.069)
	Number of observations	33,696	15,756	15,756
	Number of users	2,808	2,358	2,358
	Adjusted R^2	0.646	0.515	0.628
G4	Treatment × after	0.342***	0.112	0.040
		(0.052)	(0.070)	(0.090)
	Number of observations	31,896	14,058	14,058
	Number of users	2,658	2,161	2,161
	Adjusted R^2	0.674	0.553	0.640

Notes. User and month fixed effects are included. Robust standard errors in parentheses.
***$p < 0.01$; **$p < 0.05$; *$p < 0.1$.

Columns (2) and (3) of Table 2 present estimation results using the average log number of characters per answer and average log number of vote-ups per answer as the dependent variables. The coefficients of $Treatment_i \times After_t$ are mostly not statistically significant. Hence, the paid feature has not affected the amount of effort that hosts put into their answers or the perceived usefulness of the answers. This finding may arise because answer quality depends on contributors' innate ability or competence, which is difficult to change in a short period of time.

Our results are robust to several alternative specifications, including applying fixed effects Poisson model, using the PSM technique, and removing outliers with extreme contributions. We report the details of these robustness checks in Section C of the online appendix.

5.4. Validation

5.4.1. Falsification Test. To address the concern that the results may be driven by spurious correlations or inappropriate matching, we conduct a falsification exercise by setting the placebo cutoff dates for live hosts to six months before the real cutoff dates (Goh et al. 2015, Bapna et al. 2018) and repeating the matching and all estimations. Supposedly, the treatment effect should not be present in this exercise. If, however, our findings in Table 2 are spurious and caused by uncaptured correlated trends between user heterogeneity and contributions, then the placebo treatment may also show a statistically significant effect.

As reported in Table 3, the coefficients of $Treatment_i \times After_t$ are not statistically significant for all dependent variables, including answer quantity. This shows that our findings are specific to the paid live feature instead of other spurious trends or failure in the matching procedure.

5.4.2. Lookahead Matching. Although we have carefully matched live hosts and nonhosts using a set of observed variables, the two groups of users may still differ on some unobserved characteristics that drive their self-selection into the treatment or control group, leading to potential estimation bias (Shadish et al. 2002). To account for such unobserved user characteristics, we use the lookahead matching strategy (Hosanagar et al. 2013, Bapna et al. 2018), which matches existing hosts to users who were nonhosts at the time of matching but who would become live hosts in the future. Given that these users became hosts eventually, they should share similar unobserved characteristics governing the decision to host live talks. Specifically, we match hosts in G1 to those in G2–G4, hosts in G2 to those in G3 and G4, and hosts in G3 to those in G4.[21] We then use the same DID model to estimate the paid live feature's spillover effects.

We present the results based on the treatment and control groups under lookahead matching in Table 4. The sizes of matched samples become substantially smaller because nonhosts cannot be used for matching. Yet the estimation results are similar to those reported in Table 2 except that the coefficient of $Treatment_i \times After_t$ for answer quantity for G1 becomes statistically

Table 3. Falsification Test

CEM weighted matching		Quantity (number of answers) (1)	Quality (number of characters) (2)	Quality (number of vote-ups) (3)
G1	Treatment × after	−0.033	0.020	0.060
		(0.040)	(0.059)	(0.065)
	Number of observations	17,520	9,797	9,797
	Number of users	1,460	1,304	1,304
	Adjusted R^2	0.705	0.448	0.607
G2	Treatment × after	0.025	0.067	−0.029
		(0.036)	(0.062)	(0.068)
	Number of observations	25,608	12,728	12,728
	Number of users	2,134	1,829	1,829
	Adjusted R^2	0.701	0.496	0.667
G3	Treatment × after	0.057	0.111*	0.031
		(0.038)	(0.062)	(0.075)
	Number of observations	28,116	13,084	13,084
	Number of users	2,343	1,996	1,996
	Adjusted R^2	0.660	0.470	0.620
G4	Treatment × after	0.005	0.072	0.089
		(0.044)	(0.083)	(0.101)
	Number of observations	28,032	12,358	12,358
	Number of users	2,336	1,937	1,937
	Adjusted R^2	0.642	0.525	0.640

Notes. User and month fixed effects are included. Robust standard errors in parentheses.
$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$.

insignificant. Although the results are consistent and continue to indicate a positive spillover effect of the paid live feature on answer quantity, we interpret this result with caution as using future hosts as control users may underestimate the effect size. The future hosts could have increased their answer contributions before holding their live talks.

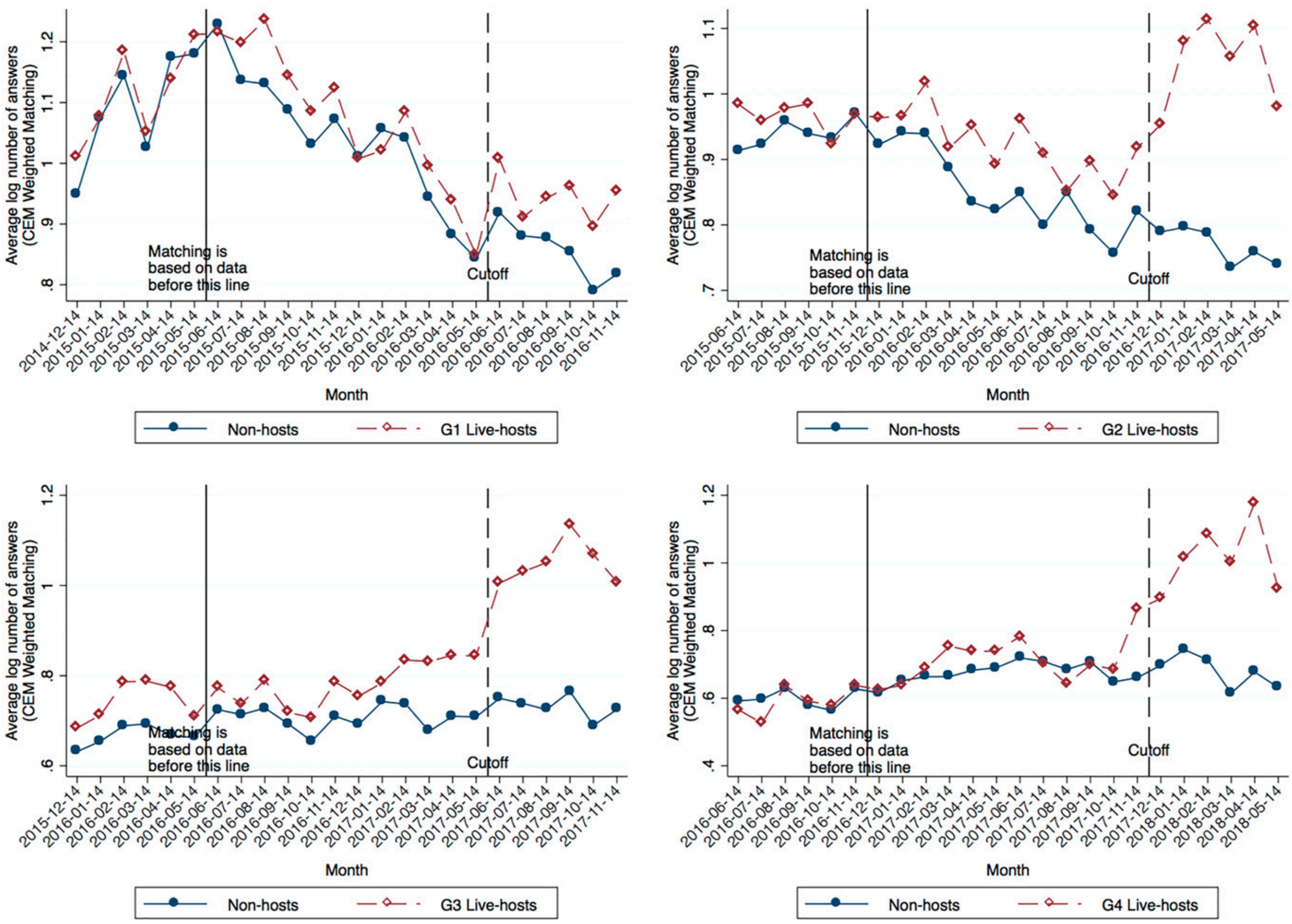
5.4.3. Earlier Matching Points. In these analyses, we set the matching points to be six months before the cutoff dates. One concern with the model-free evidence in Figure 4 is that the increase in answer contributions of treatment groups G2–G4 in the few months before the cutoff dates could have been caused by a failure in matching instead of pre–live talk reputation building. This concern is not potent as the divergence did not happen immediately after the matching, and we do not observe the same pattern for live hosts in G1 who are not supposed to act strategically before the cutoff. Nevertheless, we conduct another validation test by

Table 4. Lookahead Matching

Lookahead CEM weighted matching		Quantity (number of answers) (1)	Quality (number of characters) (2)	Quality (number of vote-ups) (3)
G1	Treatment × after	0.051	0.098	0.054
		(0.041)	(0.076)	(0.085)
	Number of observations	9,768	5,353	5,353
	Number of users	814	748	748
	Adjusted R^2	0.634	0.391	0.590
G2	Treatment × After	0.102**	−0.101	−0.012
		(0.047)	(0.097)	(0.092)
	Number of observations	8,772	4,583	4,583
	Number of users	731	656	656
	Adjusted R^2	0.615	0.398	0.611
G3	Treatment × after	0.271***	−0.068	−0.022
		(0.080)	(0.127)	(0.151)
	Number of observations	4,704	2,255	2,255
	Number of users	392	357	357
	Adjusted R^2	0.513	0.429	0.598

Notes. User and month fixed effects are included. Robust standard errors in parentheses.
$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$.

Figure 7. (Color online) Average Answer Quantity, Earlier Matching Points



Notes. The solid vertical line represents the matching point of each group. The dashed vertical line represents the cutoff date of each group.

shifting the matching points backward to one full year (instead of six months) before the cutoff dates. Longer "after matching" and "before cutoff" periods allow us to gain a better understanding of the nature of the increase near the cutoffs.

The model-free evidence on answer quantity, as presented in Figure 7, confirms that our matching is robust as the matched live hosts and nonhosts had similar trends of answer contributions for a long time after the matching points. The gap in answer contributions was present only in the few months before the cutoffs, which supports the argument that the live hosts may have strategically increased their answer contributions shortly before hosting the first talks. In addition, the fact that later live hosts did not take action until near the dates when they conducted the first live talks provides indicative evidence for the heterogeneous exposure of live hosts to the paid feature.

6. Underlying Mechanisms

6.1. Reputation Building

We next explore the reasons underlying the positive spillover of the paid feature on users' contribution to free answers. We analyze one focused explanation: that the live hosts are trying to build their reputation by contributing answers that are openly accessible by all users. A better reputation not only signals the quality and credibility of live hosts but also increases the visibility of their activities, including live talks. We measure reputation using the cumulative number of vote-ups obtained by each user, which integrates both quantity and quality of user contribution (Mamykina et al. 2011, Li et al. 2012). The cumulative number of vote-ups received by each user is displayed in a prominent position of personal achievements on the user homepage. Hence, it is easily noticeable by other Zhihu users.

We begin by exploring the relationship between the hosts' reputation before holding talks and the revenues of their first live talks. Given that both reputation and revenue are highly skewed, we perform a log–log linear regression to estimate the reputation effect. The result is reported in Table 5. Clearly, the effect of reputation on revenue is positive and statistically significant, which supports the hypothesis that higher reputation can translate into more revenue from live talks.

Table 5. Effect of Reputation on Performance of First Talk (Log–Log)

	Revenue (1)	Entrance fee (2)	Number of listeners (3)	Review rating (4)
Reputation	0.119***	0.019***	0.096***	−0.000
	(0.010)	(0.003)	(0.009)	(0.001)
Number of observations	2,024	2,024	2,024	1,928
Adjusted R^2	0.072	0.014	0.058	−0.001

***$p < 0.01$; **$p < 0.05$; *$p < 0.1$.

Specifically, doubling the number of vote-ups leads to an 8.6% increase in the revenue.

Next, we explore whether this revenue advantage is due to higher entrance fees or more listeners. Columns (2) and (3) of Table 5 show the results with entrance fee and number of listeners as the dependent variables. We find that reputation has significant positive effects on both entrance fee and audience size. Specifically, doubling the number of vote-ups is associated with a 1.3% increase in entrance fee and a 6.9% increase in number of listeners. This set of analyses indicates high-reputation hosts obtain more revenue from their talks via both higher entrance fees and larger audience sizes, and the effect of audience size is stronger than that of entrance fee. However, it is possible that high-reputation hosts gain more revenue because of better talk quality. To test this, we use the live talks' review ratings as the quality indicator and assess the effect of reputation on review rating. Column (4) of Table 5 shows that the hosts' reputation does not affect their live talks' review ratings. This suggests that the revenue premium of high-reputation hosts is caused by their better reputation instead of superior talk quality.

We now exploit users' prior reputation to corroborate the reputation-building mechanism. If the paid feature causes users to contribute more answers because of reputation building, then we expect the effect to be more pronounced for people with lower prior reputation. From Table 2, the positive spillover effect of monetary incentive is strongest for the live hosts in G4, followed by G3, G2, and G1. From Figure 4, the monthly average number of answers contributed before the cutoff is highest for users in G1, followed by G2, G3, and G4. In fact, the median cumulative number of vote-ups before the cutoff date is 19,720 for G1, but only 7,151 for G2, 1,744 for G3, and 1,020 for G4. These trends suggest that the live hosts' prior reputation is negatively correlated with the spillover effect of monetary incentive on their free answer contributions. Apparently, low-reputation live hosts made more free answer contributions after the cutoffs, probably because they wanted to build reputation via these free answer contributions.

An alternative explanation of the trends in Figure 4 is that users need time to get familiar with the live feature and appreciate the potential benefits before changing their behavior. To eliminate the influence of adoption timing, we partition the live hosts in each treatment group into two sets based on their reputation before the cutoff date. If a host's prior reputation is above the median of *all* live hosts, then the host falls into the *high reputation* set. Otherwise, the host falls into the *low reputation* set. We use the same matching procedure as explained in Section 4 to construct comparable "high reputation" and "low reputation" control group users for these two sets of live hosts and repeat the DID estimation on the matched samples.[22]

Columns (1) and (4) of Table 6 report the subgroup analysis results on answer quantity.[23] Across the four groups, the coefficients of $Treatment_i \times After_t$ for low-reputation live hosts are larger than those for high-reputation live hosts. The difference is especially noticeable for hosts who held their first talks more than one year after the launch of Zhihu Live. Specifically, the low-reputation hosts in G3 and G4 provided 35.1% and 55.6% more answers relative to low-reputation nonhosts after the cutoff date. The corresponding increases for the high-reputation hosts in G3 and G4 are 12.7% and 20.1% and less significant. Columns (2), (3), (5), and (6) of Table 6 report the results on answer quality. We do not find any pervasive effect of the live feature on answer quality for either high- or low-reputation hosts.

These results, that the spillover effect on answer quantity is stronger for low-reputation hosts in percentage terms, can also be due to the low starting basis of low-reputation hosts. To further validate the results, we examine the effect in absolute terms using the unlogged number of answers. The estimates, as reported in Table H4 of the online appendix, show that the size of the spillover effect is larger for low- than for high-reputation hosts in terms of both magnitude and significance.

Taken together, we find strong evidence that live hosts with lower reputation are more motivated to contribute free answers after the cutoff date. This is consistent with reputation building, by which live hosts raise their visibility and cultivate potential listeners by offering more free answers.

6.2. Alternative Mechanisms

We now consider two compelling alternative mechanisms that could give rise to a positive spillover effect of monetary incentives on free knowledge contribution.

Table 6. Subgroup Analysis Based on Prior Reputation

		High reputation			Low reputation		
CEM weighted matching		Quantity (number of answers) (1)	Quality (number of characters) (2)	Quality (number of vote-ups) (3)	Quantity (number of answers) (4)	Quality (number of characters) (5)	Quality (number of vote-ups) (6)
G1	Treatment × after	0.092*	0.102	−0.115	0.160***	0.295**	0.453***
		(0.053)	(0.079)	(0.099)	(0.062)	(0.132)	(0.143)
	Number of observations	8,436	5,121	5,121	7,272	3,365	3,365
	Number of users	703	644	644	606	522	522
	Adjusted R^2	0.698	0.460	0.575	0.605	0.429	0.514
G2	Treatment × after	0.141***	0.086	0.086	0.183***	−0.058	−0.093
		(0.052)	(0.071)	(0.087)	(0.050)	(0.111)	(0.115)
	Number of observations	9,396	5,421	5,421	11,364	4,998	4,998
	Number of users	783	719	719	947	806	806
	Adjusted R^2	0.713	0.515	0.630	0.581	0.450	0.519
G3	Treatment × after	0.120*	0.002	0.089	0.301***	0.127	0.029
		(0.064)	(0.071)	(0.096)	(0.056)	(0.097)	(0.114)
	Number of observations	9,456	5,073	5,073	14,232	6,012	6,012
	Number of users	788	676	676	1,186	993	993
	Adjusted R^2	0.725	0.507	0.585	0.565	0.459	0.528
G4	Treatment × after	0.183*	0.100	0.138	0.442***	0.192*	0.052
		(0.097)	(0.100)	(0.154)	(0.063)	(0.101)	(0.122)
	Number of observations	8,436	4,243	4,243	12,492	5,000	5,000
	Number of users	703	601	601	1,041	824	824
	Adjusted R^2	0.713	0.582	0.639	0.620	0.493	0.460

Notes. User and month fixed effects are included. Robust standard errors in parentheses.
$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$.

6.2.1. Explicit Promotion. One explanation of our findings is that live hosts use free answers as an explicit promotion channel. The typical way to do this is to provide a short description and a hyperlink to the promoted talk. When other users see the promotions in the answers, they may learn about the live talk and pay to listen. To identify such contributions, we check whether the answers contain a URL with the string "zhihu.com/lives/". In our sample, 7,343 answers contain URLs of live talks out of a total of 187,013 answers provided by live hosts after the paid feature was introduced. The low proportion of answers with live talk URLs suggests that explicit promotion is not likely to be the driving force of the positive spillover effect.

To ensure this alternative explanation is not the primary cause of our findings, we exclude all answers containing explicit promotion from the data and repeat the estimation.[24] Table 7 reports the results. The findings are similar to those reported in Table 2 except that the effect size has decreased slightly, which is expected because removing the answers with explicit promotion inevitably reduces the postcutoff contributions of live hosts to free answers.

6.2.2. Reciprocity. The median total revenue of live hosts from hosting live talks is around US$1,500 in our sample. With such a high reward, the live hosts may repay the favor by undertaking actions helpful to the community, such as providing more free answers. In an online review platform, Pu et al. (2017) show that reviewers indeed provide more and longer reviews after they begin to receive free products from the platform. They attribute this behavioral change to reciprocity.

Intuitively, if the positive spillover is driven by reciprocity, we should expect the effect to be more pronounced for high-reputation hosts as they obtain more revenue from holding live talks (see Table 5). However, as reported in Section 6.1, low-reputation hosts are more motivated to contribute free answers, indicating reciprocity is unlikely to drive our findings.

To further separate the effect of reputation building from positive reciprocity, for each live host, we drop all observations starting from the month of the host's first talk and redo all analysis.[25] As the live hosts do not receive any money before their first live talks, any change in their behavior should not be caused by the motivation to reciprocate. By contrast, reputation building could happen before their first live talks as shown in Figures 4 and 7 and discussed in Section 5.4.3. Table 8 reports the estimation results based on this pruned data set. Once again, the results are consistent with those reported in Table 2, suggesting that the positive spillover effect persists without reciprocity.

Table 7. Excluding Explicit Ads

CEM weighted matching		Quantity (number of answers) (1)	Quality (number of characters) (2)	Quality (number of vote-ups) (3)
G1	Treatment × after	0.079**	0.084	−0.020
		(0.040)	(0.055)	(0.067)
	Number of observations	22,380	11,973	11,973
	Number of users	1,865	1,645	1,645
	Adjusted R^2	0.704	0.482	0.628
G2	Treatment × after	0.145***	−0.000	−0.022
		(0.033)	(0.054)	(0.060)
	Number of observations	29,916	14,621	14,621
	Number of users	2,493	2,160	2,160
	Adjusted R^2	0.708	0.490	0.668
G3	Treatment × after	0.197***	0.023	0.061
		(0.040)	(0.058)	(0.070)
	Number of observations	33,696	15,698	15,698
	Number of users	2,808	2,354	2,354
	Adjusted R^2	0.647	0.511	0.626
G4	Treatment × after	0.298***	0.046	0.005
		(0.051)	(0.072)	(0.091)
	Number of observations	31,764	13,953	13,953
	Number of users	2,647	2,152	2,152
	Adjusted R^2	0.675	0.530	0.627

Notes. User and month fixed effects are included. Robust standard errors in parentheses.
$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$.

7. Long-Term Effect

So far, we have been focusing on the paid feature's short-term effect by examining the live hosts' answer contribution in the six months before and after the cutoff date. It is not clear how live hosts would behave in the long run. Our long panel, which covers two years of observations after the launch of Zhihu Live, provides us with an opportunity to investigate the long-term behavior of live hosts who held their first talks at a relatively earlier time (i.e., G1–G3). Specifically, for each treatment group, we include the entire period after the cutoff date and divide it into multiple periods of six months each. The new model specifications for G1–G3 are

$$\begin{aligned} y_{it} = \beta_0 &+ \beta_1 Treatment_i \times After\_1_t \\ &+ \beta_2 Treatment_i \times After\_2_t \\ &+ \beta_3 Treatment_i \times After\_3_t \\ &+ \beta_4 Treatment_i \times After\_4_t + u_i + \tau_t + \epsilon_{it}, \end{aligned} \tag{2}$$

$$\begin{aligned} y_{it} = \beta_0 &+ \beta_1 Treatment_i \times After\_1_t \\ &+ \beta_2 Treatment_i \times After\_2_t \\ &+ \beta_3 Treatment_i \times After\_3_t + u_i + \tau_t + \epsilon_{it}, \end{aligned} \tag{3}$$

$$\begin{aligned} y_{it} = \beta_0 &+ \beta_1 Treatment_i \times After\_1_t \\ &+ \beta_2 Treatment_i \times After\_2_t + u_i + \tau_t + \epsilon_{it}, \end{aligned} \tag{4}$$

where $After\_1_t$ equals one in the first six-month period after the cutoff date, $After\_2_t$ equals one in the second six-month period after the cutoff date, and so on. With this specification, the coefficient β_1 is the original (short-term) DID estimator. The coefficients β_2, β_3, and β_4 capture separate long-term effects. Note that we cannot examine the long-term effect of live hosts in G4 because we only have six months of observations for them after the cutoff date.

Column (1) of Table 9 presents the estimation results on answer quantity. We do not observe a long-term effect of the paid feature with the exception of G1 in which a positive spillover effect still existed in the second six-month period after the cutoff date. Columns (2) and (3) of Table 9 suggest that there is no long-term effect on answer quality either.

These estimations pool all live hosts in the same group without considering the heterogeneity within the group. Next, we explore whether the long-term effect varies across different types of hosts.

7.1. Long-Term Effect for Low- and High-Reputation Hosts

In Section 6.1, we show that low-reputation hosts have a greater incentive than high-reputation hosts to boost their reputation by contributing free answers in the short term after the cutoff date. Now, we examine how these two types of hosts behave in the long run with respect to their free answer contributions.

Columns (1) and (4) of Table 10 present the long-term results on answer quantity for high- and low-reputation hosts. We do not observe a long-term effect

Table 8. Excluding Observations Starting from the Month of First Talk

CEM weighted matching		Quantity (number of answers) (1)	Quality (number of characters) (2)	Quality (number of vote-ups) (3)
G1	Treatment × after	0.074*	0.097	0.004
		(0.043)	(0.066)	(0.079)
	Number of observations	21,270	11,297	11,297
	Number of users	1,865	1,634	1,634
	Adjusted R^2	0.711	0.487	0.631
G2	Treatment × after	0.166***	−0.002	0.066
		(0.039)	(0.081)	(0.084)
	Number of observations	28,374	13,667	13,667
	Number of users	2,493	2,141	2,141
	Adjusted R^2	0.714	0.492	0.667
G3	Treatment × after	0.185***	0.100	0.087
		(0.049)	(0.082)	(0.090)
	Number of observations	32,165	14,821	14,821
	Number of users	2,808	2,334	2,334
	Adjusted R^2	0.655	0.517	0.625
G4	Treatment × after	0.292***	0.222**	0.127
		(0.056)	(0.093)	(0.113)
	Number of observations	31,125	13,573	13,573
	Number of users	2,658	2,149	2,149
	Adjusted R^2	0.678	0.556	0.640

Notes. User and month fixed effects are included. Robust standard errors in parentheses.
***$p < 0.01$; **$p < 0.05$; *$p < 0.1$.

of the paid feature on either high- or low-reputation hosts except for the low-reputation hosts in G1 for which a positive spillover effect still existed in the second six-month period after the cutoff date. Same as the analysis in Section 6.1, the short-term increase in the quantity of free answer contributions is mostly made by low-reputation hosts. Columns (2), (3), (5), and (6) of Table 10 indicate that there is no pervasive long-term effect on answer quality for either group.

This set of results suggests the reputation-building efforts of both high- and low-reputation hosts did not persist in the long run despite the heterogeneity in their short-term efforts. Although low-reputation hosts made more free answer contributions in the short term, the increase in their free contributions was not sustained over the long term.

7.2. Long-Term Effect for Short- and Long-Lived Hosts

So far, the evidence suggests live hosts made more free answer contributions, but this effort is not persistent. Could this be due to attrition, that is, their interest in giving live talks has dissipated over time? Live hosts in our sample have different life spans in holding talks. Intuitively, hosts who held talks within a short period and stopped afterward may behave quite differently from those who continued for a longer period. To investigate this possibility, we partition the live hosts in each group into two sets. If a host only held talks in that six-month period, then we put the host into the *short-lived* set. Otherwise, we put the host into the *long-lived* set. Here again, we randomly assign half of the nonhosts to match the short-lived set and the other half to match the long-lived set to generate comparable control group users.

Columns (1)–(3) of Table 11 report the long-term results for short-lived hosts. The coefficients of $Treatment_i \times After\_3_t$ and $Treatment_i \times After\_4_t$ for answer quantity are negative and statistically significant. The coefficients of $Treatment_i \times After\_2_t$ for answer quantity, although not statistically significant, are negative across all three groups. We also observe a significant negative effect of the paid feature on the long-term quality of answers in terms of the number of vote-ups received. This set of results suggests that the introduction of monetary incentives may hurt the free knowledge contribution of users after they stop holding live talks, which is consistent with the crowding-out theory (Titmuss 1970, Frey and Oberholzer-Gee 1997, Deci et al. 1999). Monetary incentives may crowd out individuals' intrinsic and image motivations. The overall motivation may fall below the preintervention level after the incentive is removed. In Section D of the online appendix, we address two other explanations that may also drive the observed negative long-run spillover effect for short-lived hosts.

We next consider the long-term results for long-lived hosts. Column (4) of Table 11 shows that the coefficients of $Treatment_i \times After\_1_t$ and $Treatment_i \times After\_2_t$ are positive and statistically significant across

Table 9. Long-Term Effect for All Hosts

		All Hosts		
CEM weighted matching		Quantity (number of answers) (1)	Quality (number of characters) (2)	Quality (number of vote-ups) (3)
G1	Treatment × after_1	0.090**	0.110**	0.016
		(0.040)	(0.055)	(0.067)
	Treatment × after_2	0.094**	0.052	−0.013
		(0.047)	(0.064)	(0.075)
	Treatment × after_3	−0.005	−0.016	−0.041
		(0.050)	(0.070)	(0.088)
	Treatment × after_4	−0.078	−0.028	−0.111
		(0.053)	(0.075)	(0.097)
	Number of observations	55,950	27,595	27,595
	Number of users	1,865	1,736	1,736
	Adjusted R^2	0.665	0.448	0.593
G2	Treatment × after_1	0.171***	0.035	0.032
		(0.033)	(0.053)	(0.059)
	Treatment × after_2	0.035	−0.009	−0.045
		(0.040)	(0.059)	(0.068)
	Treatment × after_3	−0.070	0.004	−0.123
		(0.044)	(0.069)	(0.080)
	Number of observations	59,832	28,459	28,459
	Number of users	2,493	2,269	2,269
	Adjusted R^2	0.664	0.474	0.636
G3	Treatment × after_1	0.239***	0.065	0.066
		(0.041)	(0.055)	(0.069)
	Treatment × after_2	0.066	−0.103*	−0.116
		(0.043)	(0.057)	(0.082)
	Number of observations	50,544	23,222	23,222
	Number of users	2,808	2,437	2,437
	Adjusted R^2	0.627	0.509	0.610

Notes. User and month fixed effects are included. Robust standard errors in parentheses.
*** $p < 0.01$; ** $p < 0.05$; * $p < 0.1$.

all three treatment groups. The coefficients for the longer periods are not statistically significant. Therefore, it seems that the positive spillover effect of the paid feature on answer quantity extends beyond the short term, but the effect size decays over time. By examining our data, we find that although long-lived hosts still held talks after the first six-month period, many of them eventually stopped in longer periods, leading to a reduction in their free contributions over time. In columns (5) and (6) of Table 11, we do not observe a pervasive long-term effect on answer quality for long-lived hosts.

To explore why some live hosts stopped holding talks after a short period, whereas some other live hosts continued, we compare the first talks of short- and long-lived hosts in terms of entrance fee, number of listeners, total revenue, number of reviews, and average review rating. Table H5 in the online appendix reports this result. Evidently, the long-lived hosts' first talks were more successful than the short-lived hosts' first talks. Note that, even though, on average, high-reputation hosts generate more revenue from their talks than low-reputation hosts, there was a great disparity between low-reputation and short-lived hosts and between high-reputation and long-lived hosts.[26] The converging finding is that the paid feature has significantly encouraged live hosts to make more free answer contributions, but this effect is heterogeneous and not persistent. It could even lead to crowding out of short-lived hosts in terms of free answer contributions.

8. Discussion and Conclusions

To our knowledge, this is among the first empirical studies on the spillover effect of monetary incentives for a rewarded activity on users' contribution to another nonrewarded activity on an online knowledge platform. We study live hosts' behavioral changes in their free answer contributions in both quantity and quality by exploiting a natural experiment on Zhihu and applying CEM matching and DID techniques. We find that the paid feature caused live hosts to contribute more answers without sacrificing the average quality in the short run (cf. Kuang et al. 2019). The increase in free answer contribution is stronger among low-reputation hosts, and reputation is correlated

Table 10. Long-Term Effect for High-Reputation vs. Low-Reputation Hosts

		High reputation			Low reputation		
CEM weighted matching		Quantity (number of answers) (1)	Quality (number of characters) (2)	Quality (number of vote-ups) (3)	Quantity (number of answers) (4)	Quality (number of characters) (5)	Quality (number of vote-ups) (6)
G1	Treatment × after_1	0.092*	0.095	−0.105	0.160***	0.294**	0.500***
		(0.053)	(0.081)	(0.098)	(0.062)	(0.132)	(0.142)
	Treatment × after_2	0.057	−0.023	−0.120	0.172**	0.187	0.418***
		(0.070)	(0.092)	(0.102)	(0.083)	(0.153)	(0.154)
	Treatment × after_3	0.004	−0.051	−0.100	0.037	0.050	0.167
		(0.073)	(0.104)	(0.120)	(0.070)	(0.158)	(0.188)
	Treatment × after_4	−0.085	−0.102	−0.243*	−0.000	0.088	0.277
		(0.079)	(0.108)	(0.128)	(0.079)	(0.173)	(0.210)
	Number of observations	21,090	11,835	11,835	18,180	7,662	7,662
	Number of users	703	667	667	606	557	557
	Adjusted R^2	0.669	0.424	0.539	0.596	0.386	0.515
G2	Treatment × after_1	0.141***	0.102	0.094	0.183***	−0.042	0.021
		(0.052)	(0.071)	(0.085)	(0.050)	(0.105)	(0.113)
	Treatment × after_2	0.026	−0.069	−0.033	0.050	0.108	0.045
		(0.065)	(0.076)	(0.093)	(0.058)	(0.116)	(0.119)
	Treatment × after_3	−0.078	−0.034	−0.059	−0.080	0.061	−0.099
		(0.070)	(0.098)	(0.106)	(0.063)	(0.132)	(0.163)
	Number of observations	18,792	10,401	10,401	22,728	9,678	9,678
	Number of users	783	733	733	947	855	855
	Adjusted R^2	0.697	0.480	0.601	0.546	0.440	0.513
G3	Treatment × after_1	0.120*	0.020	0.081	0.301***	0.124	0.030
		(0.064)	(0.071)	(0.096)	(0.056)	(0.095)	(0.113)
	Treatment × after_2	0.017	−0.104	0.027	0.086	−0.049	−0.148
		(0.079)	(0.071)	(0.122)	(0.058)	(0.097)	(0.129)
	Number of observations	14,184	7,374	7,374	21,348	8,886	8,886
	Number of users	788	688	688	1,186	1,032	1,032
	Adjusted R^2	0.715	0.497	0.573	0.549	0.461	0.529

Notes. User and month fixed effects are included. Robust standard errors in parentheses.
***$p < 0.01$; **$p < 0.05$; *$p < 0.1$.

with live talk revenues. Together with evidence that the positive spillover exists after removing explicit promotion and reciprocity, we believe reputation building is a likely mechanism leading to the increase in free answer contribution.

We do not observe a long-term effect of the paid feature on either high- or low-reputation hosts. However, the free contributions of hosts in the long run differ markedly depending on their life spans. Specifically, we find that the positive spillover effect decays in the long run among long-lived hosts. The long-term effect is even negative among short-lived hosts. Although the incentive to build reputation did motivate live hosts to contribute more free answers in the short run, the effect is heterogeneous and, apparently, lasts longer for long-lived hosts. We interpret the long-run negative spillover for short-lived hosts as the crowding-out effect because it exists among hosts who continue to stay in the platform (despite the fact that they no longer hold live talks) and applies to all short-lived hosts regardless of their ability to generate live talk revenues out of their prevailing reputation. The long-run negative spillover is not observed for long-lived hosts because the reputation building incentive is still at work as long as they continue to hold live talks.

We now discuss the overall net impact of the paid feature over the whole period of two years. At the platform level, there was a reduction of 32.8% in the monthly contribution of nonhosts in the two years after the policy, whereas the corresponding reduction for live hosts was 18.3%. It seems that although the answer contribution of users on Zhihu followed a downward trend in our study period, the introduction of the live feature has alleviated this reduction by motivating hosts to contribute more free answers. The pooled analysis using a single cutoff date detailed in Section E of the online appendix yields consistent results.[27] Therefore, we conclude that the introduction of monetary incentives has a positive overall effect on free knowledge contributions in our study period. We acknowledge, however, that it is difficult to predict how the spillover effect would evolve in the longer future.

The findings from our research offer important implications for owners of knowledge platforms. To

Table 11. Long-Term Effect for Short-Lived vs. Long-Lived Hosts

		Short lived			Long lived		
CEM weighted matching		Quantity (number of answers) (1)	Quality (number of characters) (2)	Quality (number of vote-ups) (3)	Quantity (number of answers) (4)	Quality (number of characters) (5)	Quality (number of vote-ups) (6)
G1	Treatment × after_1	0.041	0.125	−0.000	0.137**	0.160**	0.099
		(0.054)	(0.097)	(0.112)	(0.059)	(0.073)	(0.092)
	Treatment × after_2	−0.041	−0.115	−0.142	0.186***	0.212***	0.124
		(0.067)	(0.108)	(0.120)	(0.066)	(0.079)	(0.109)
	Treatment × after_3	−0.112*	−0.048	−0.169	0.046	0.069	0.107
		(0.067)	(0.114)	(0.144)	(0.073)	(0.085)	(0.113)
	Treatment × after_4	−0.196***	−0.199	−0.369**	−0.031	0.187*	0.127
		(0.070)	(0.128)	(0.152)	(0.078)	(0.104)	(0.140)
	Number of observations	22,380	10,649	10,649	24,090	12,462	12,462
	Number of users	746	692	692	803	757	757
	Adjusted R^2	0.657	0.446	0.580	0.689	0.428	0.581
G2	Treatment × after_1	0.097**	−0.005	0.001	0.281***	0.073	0.065
		(0.040)	(0.070)	(0.081)	(0.062)	(0.088)	(0.098)
	Treatment × after_2	−0.061	−0.015	−0.147	0.154**	0.035	0.057
		(0.048)	(0.078)	(0.094)	(0.076)	(0.097)	(0.115)
	Treatment × after_3	−0.164***	0.058	−0.206*	0.034	0.008	−0.063
		(0.052)	(0.086)	(0.107)	(0.084)	(0.124)	(0.142)
	Number of observations	28,992	13,407	13,407	23,232	11,375	11,375
	Number of users	1,208	1,090	1,090	968	898	898
	Adjusted R^2	0.684	0.480	0.630	0.634	0.432	0.638
G3	Treatment × after_1	0.161***	0.090	0.015	0.383***	0.047	0.131
		(0.048)	(0.077)	(0.095)	(0.083)	(0.084)	(0.106)
	Treatment × after_2	−0.032	−0.189**	−0.329***	0.227***	−0.016	0.118
		(0.052)	(0.081)	(0.110)	(0.085)	(0.085)	(0.133)
	Number of observations	25,182	11,190	11,190	18,666	9,096	9,096
	Number of users	1,399	1,205	1,205	1,037	928	928
	Adjusted R^2	0.618	0.507	0.594	0.635	0.530	0.649

Notes. User and month fixed effects are included. Robust standard errors in parentheses.
***$p < 0.01$; **$p < 0.05$; *$p < 0.1$.

encourage the contribution of more and higher quality content, many platforms have started to introduce monetary rewards for providing premium content, a practice that leads to a new knowledge-sharing economy. For knowledge platforms with multiple types of contributions, paying for one specific type of activity may affect users' contributions to other nonrewarded activities. The success of Zhihu Live in facilitating users' contributions in the focal activity of hosting live talks and the free activity of answering questions seems to suggest that this can be a viable business model for Q&A communities that have been struggling to sustain user contributions and make profits. When the payment for one knowledge activity is in force, platform owners may consider introducing other related nonrewarded activities to exploit users' motivation to boost their reputation and win a paying audience.

Despite the positive findings in the short run, the evidence on long-term spillover is mixed. If users stop contributing to the paid activity, perhaps because they lose interest, their contribution to nonrewarded activities could fall below the preintervention period, eventually causing a negative spillover effect. If users continue engaging in the paid activity to receive the monetary incentives, the positive spillover effect may last longer. Our long-term effect analysis revitalizes the debate of whether extrinsic incentives crowd out users' intrinsic motivation to make voluntary contributions. It is plausible that monetary incentives encourage more contributions from some people but crowd out others' contributions. Platform owners must explicitly account for (possibly endogenous) user heterogeneity when deliberating whether to introduce monetary incentives to stimulate knowledge contribution. They should preempt users' behaviors in the long run and design the response strategies accordingly. For example, they could organize regular workshops to educate users or designate specific staff to address users' concerns so that users are more likely to continue engaging in rewarded activities. In addition, platform owners could consider providing extra rewards tied to users' continuous engagement in rewarded activities to induce a more sustained positive spillover effect.

Our analysis and findings hint at an interesting dynamic interaction between monetary incentives and users' knowledge contribution. Previous studies on the direct effect of monetary incentives have shown that the amount of compensation is positively associated with the performance in the rewarded behavior (e.g., Gneezy et al. 2011). Here, the monetary incentive itself is endogenous. The reward amount depends on the size of the audience, which, as we show in Table 5, is related to the host's reputation. Live hosts might contribute more free answers to build their reputation in the hope that a better reputation will help bring more monetary returns for their live talks. Future research should explore paid contributors' response functions when the reward is dependent on other characteristics.

Note that our setting is a two-sided market in which knowledge providers charge consumers for access to specific knowledge. Similar to other two-sided platforms, there is asymmetric information in that the quality of the paid knowledge is unobservable by consumers before payment. A direct communication channel between knowledge providers and consumers does not exist either. Under these circumstances, knowledge providers may make free and publicly accessible contributions to signal their quality and raise public awareness. We believe our results are generalizable to platforms featuring complementary goods when one good serving as a quality indicator is accessible by consumers for free, whereas the other good is chargeable. Such platforms are common on the internet, for example, MOOC platforms, consumer-to-consumer marketplaces, and software download sites.

Our study is subject to several limitations. First, we define nonhosts as users who have not hosted any live talk by the end of the sampling period. Some nonhosts could well be building their reputation for future live talks by answering more questions in our sample. This tends to bias our findings downward. The real spillover effect could be more positive if we have misclassified some future live hosts as nonhosts. Future research should explore exogenous changes in live host propensity, for example, when some users cannot benefit from the monetary incentive because of prespecified restrictions.

Second, nonhosts may be demotivated and reduce their voluntary contributions because of the paid feature. If this happens, the DID method may not unambiguously identify the live hosts' responses. This is a general limitation of DID and natural experiments when all subjects are aware of the treatments. We believe the fact that the nonhosts' answer contributions followed the original trend after the cutoff dates, as presented in Figures 4 and 7, eases this concern. Nevertheless, we cannot rule out the possibility that the motivation of nonhosts may change in the future.

Third, we find a negative effect of monetary incentive on free answer contributions of short-lived hosts in the long term. We attribute it to the crowding out of intrinsic and image motivations. However, there could be a selection bias in that hosts make the decision to stop holding talks themselves. A randomized control trial design would alleviate this concern to a large extent.

Fourth, we measure answer quality by number of characters and vote-ups. Although both of these measures have been used in previous studies as quality indicators (e.g., Harper et al. 2008, Chen et al. 2010), neither is perfect. To further validate the findings on answer quality, we hired human workers from a large crowdsourcing platform to manually rate the quality of a subset of answers and redo the analysis using this subset of data. The detailed procedure and results are presented in Section F of the online appendix. The results with human ratings continue to suggest that the paid feature did not result in any significant change in answer quality.

Another limitation is that we cannot separate the monetary incentive effect from the salience effect. The spillover could be caused by the mere presence of Zhihu Live (a new service) instead of the monetary reward, in which case the live hosts may set low or insignificant entrance fees. However, only 3.1% of the live talks in our data charge below the default rate of 9.99 RMB.[28] On the contrary, 59.3% of the talks charge above this rate, suggesting that the live hosts are motivated by financial returns. An additional test in Section G of the online appendix shows that the positive spillover effect is much larger for live hosts who charge higher entrance fees, indicating that monetary incentive is a major driver of the knowledge spillover. Despite this supporting evidence, it would be meaningful if future research can directly tease out the salience effect, perhaps by using a controlled experiment.

We conclude this paper by suggesting several future research extensions. First, we can offer incentives in multiple ways, for example, sharing advertising revenue, letting consumers tip contributors, or letting contributors charge access fees. It will be meaningful to compare different payment schemes in terms of the direction and magnitude of the spillover they generate. Second, we study only one knowledge-sharing platform. Introducing a monetary incentive on one platform may create spillover on other platforms. Expanding this study to include other knowledge-sharing platforms accessible by the same users would help paint a more complete picture of the spillover effect resulting from the monetary incentives. Finally, qualitative research might help explicate live hosts' intentions. Do they contribute more free knowledge because of quality signaling or simply raising more awareness? Why do they stop holding live talks and reduce their contributions afterward? Answering questions such as these may help platform owners design better incentive schemes to facilitate knowledge contribution in online communities.

Acknowledgments
The authors thank the senior editor, associate editor, and three anonymous reviewers for their detailed and constructive comments during the review process. The authors also thank the seminar participants at Fudan University, Hong Kong University of Science and Technology, Nanyang Technological University, University of Auckland, the 28th Workshop on Information Systems and Economics (WISE 2017), and the 22nd Pacific Asia Conference on Information Systems (PACIS 2018) Doctoral Consortium for their valuable feedback.

Endnotes

[1] We obtained the ratio from "How Youtube Ad Revenue Works" by Eric Rosenberg on Investopedia. See http://www.investopedia.com/articles/personal-finance/032615/how-youtube-ad-revenue-works.asp (accessed February 2021).

[2] This information is based on the official FAQ of WordPress.com. See https://wordads.co/faq/ (accessed February 2021).

[3] As of February 2021, about 70,000 instructors on Udemy have offered more than 155,000 courses that have attracted more than 40 million students worldwide. See https://ufbsupport.udemy.com/hc/en-us/articles/115010907827-Team-Plan-FAQs (accessed February 2021).

[4] As of December 2017, Zhihu Live has hosted about 7,000 live talks generating revenue of more than US$29.7 million from five million paid listeners. See https://zhuanlan.zhihu.com/p/33135547 (accessed February 2021).

[5] Interestingly, a question in Zhihu asked, "Would users still be willing to answer questions for free after the launch of Zhihu Live?" This indicates that the community users are concerned about free knowledge contribution after the new service. See https://www.zhihu.com/question/58327192 (accessed February 2021).

[6] We define the cutoff dates based on the six-month period in which the live hosts held their first talks. Please refer to Section 4 for more details.

[7] The term "reputation building" here encompasses both signaling quality and raising public awareness.

[8] Image motivation captures "the desire to be liked and respected by others and by one's self" (Ariely et al. 2009).

[9] We obtained the numbers from 36Kr. See https://www.36kr.com/p/1025344651233288 (accessed February 2021).

[10] Zhihu makes the approval decision based on a number of factors, such as the completeness of personal information and related qualification certificates, and whether the outline of the talk is sufficiently detailed. It does not consider users' historical activities on Zhihu.

[11] An alternative sampling strategy is to collect the users' followers instead of followees. However, the set of users obtained using such a strategy is much less comparable to the live hosts. We thank the associate editor for suggesting our current sampling approach.

[12] Starting from July 2016, Zhihu allows qualified organizations, such as private companies, nonprofit institutions, and government agencies, to own organization accounts.

[13] Both answering questions and writing articles can be considered as free knowledge contribution. However, article contribution is relatively rare and, more importantly, before March 2016, Zhihu allowed only invited users to write articles. Therefore, we do not focus on article contribution in this paper. Additional analysis in Section A of the online appendix shows that there is a positive spillover effect on the quantity of article contribution, but this effect is mainly driven by the need to promote live talks.

[14] We follow Chen et al. (2010) and Burtch et al. (2017) in using the number of characters and vote-ups from other users to measure contribution quality. Note that both of these measures are undefined (missing) if user i did not answer any question in month t.

[15] We add one to the variables to avoid taking logarithm of zero.

[16] Although this design helps anchor the time when the live hosts were exposed to the live talk feature, it may suffer from the fact that live hosts may have changed their answer contributions before holding the first talks, which could have contributed to an increase in the "before" period contributions. This is corroborated by the analysis in Section 5.4.3. Therefore, the magnitude of the effect might be underestimated. Nevertheless, for live hosts who conducted the first talks in the later part of the six-month period, the anticipatory answer contributions are included in the "after" period. An alternative approach to the staged design is to use individual cutoff dates specific to each host with a predefined number of anticipation periods. Please refer to Section B of the online appendix for a detailed explanation of our choice of the staged design and the similar estimation results obtained using individual cutoffs.

[17] We coarsen the first variable using eight equally spaced cut points and the other variables using six equally spaced cut points. The balance check of variables before and after matching is presented in Table H2 of the online appendix. Evidently, the hosts and nonhosts are much more comparable after the matching.

[18] We checked the daily search index for the keyword "Zhihu Live" on Baidu, the most popular search engine in China, and found that users started to search for this keyword only after the paid feature introduction date.

[19] We calculate effect size by $\exp(0.090) - 1 = 9.4\%$. The same calculation applies to all effect size estimates reported below.

[20] Kuang et al. (2019) examine four months before and after the launch of Zhihu Live and do not find any spillover effect of monetary incentive on answer quantity. Our analysis, accounting for heterogeneous exposure of users to the "live host" treatment, shows that the introduction of the paid feature does create a positive spillover effect on answer quantity.

[21] G4 comprises users who became hosts in the last six-month period in our data set. Therefore, we cannot perform lookahead matching for G4.

[22] We randomly assign half of the nonhosts to match the high reputation set and the other half to match the low reputation set.

[23] Table H3 in the online appendix presents an alternative specification with the low reputation variable across both the treatment and control groups as a moderation variable. The results are consistent with the subgroup analysis reported in Table 6.

[24] We also exclude 13 answers containing URLs of live talks mentioned by nonhosts.

[25] The control group remains unchanged as nonhosts did not hold any live talks.

[26] In our sample, 65.1% of the low-reputation hosts and 50.6% of the high-reputation hosts were short-lived, and 34.9% of the low-reputation hosts and 49.4% of the high-reputation hosts were long-lived.

[27] Note that the estimate is conservative as later live hosts may not be aware of the policy in the earlier periods.

[28] Zhihu no longer allows users to set an entrance fee below 9.99 RMB. This option was available during our sample period.

References

Ariely D, Bracha A, Meier S (2009) Doing good or doing well? Image motivation and monetary incentives in behaving prosocially. *Amer. Econom. Rev.* 99(1):544–555.

Bapna R, Ramaprasad J, Umyarov A (2018) Monetizing freemium communities: Does paying for premium increase social engagement? *Management Inform. Systems Quart.* 42(3):719–735.

Barnett ML, Jermier JM, Lafferty BA (2006) Corporate reputation: The definitional landscape. *Corporate Reputation Rev.* 9(1):26–38.

Bénabou R, Tirole J (2006) Incentives and prosocial behavior. *Amer. Econom. Rev.* 96(5):1652–1678.

Blackwell M, Iacus SM, King G, Porro G (2009) CEM: Coarsened exact matching in Stata. *Stata J.* 9(4):524–546.

Brüggen A, Moers F (2007) The role of financial incentives and social incentives in multi-task settings. *J. Management Accounting Res.* 19(1):25–50.

Burtch G, Hong Y, Bapna R, Griskevicius V (2017) Stimulating online reviews by combining financial incentives and social norms. *Management Sci.* 64(5):2065–2082.

Charness G, Gneezy U (2009) Incentives to exercise. *Econometrica* 77(3):909–931.

Chen Y, Ho TH, Kim YM (2010) Knowledge market design: A field experiment at Google Answers. *J. Public Econom. Theory* 12(4):641–664.

Chen W, Wei X, Zhu KX (2018) Engaging voluntary contributions in online communities: A hidden Markov model. *Management Inform. Systems Quart.* 42(1):83–100.

Deci EL (1971) Effects of externally mediated rewards on intrinsic motivation. *J. Personality Soc. Psych.* 18(1):105–115.

Deci EL, Koestner R, Ryan RM (1999) A meta-analytic review of experiments examining the effects of extrinsic rewards on intrinsic motivation. *Psych. Bull.* 125(6):627–668.

Falk A, Fischbacher U (2006) A theory of reciprocity. *Games Econom. Behav.* 54(2):293–315.

Fershtman C, Gandal N (2007) Open source software: Motivation and restrictive licensing. *Internat. Econom. Econom. Policy* 4(2):209–225.

Frey BS, Oberholzer-Gee F (1997) The cost of price incentives: An empirical analysis of motivation crowding-out. *Amer. Econom. Rev.* 87(4):746–755.

Gneezy U, Rustichini A (2000) Pay enough or don't pay at all. *Quart. J. Econom.* 115(3):791–810.

Gneezy U, Meier S, Rey-Biel P (2011) When and why incentives (don't) work to modify behavior. *J. Econom. Perspect.* 25(4):191–209.

Goes PB, Guo C, Lin M (2016) Do incentive hierarchies induce user effort? Evidence from an online knowledge exchange. *Inform. Systems Res.* 27(3):497–516.

Goh KY, Hui KL, Png IP (2015) Privacy and marketing externalities: Evidence from do not call. *Management Sci.* 61(12):2982–3000.

Gouldner AW (1960) The norm of reciprocity: A preliminary statement. *Amer. Sociol. Rev.* 25(2):161–178.

Hann IH, Roberts JA, Slaughter SA (2013) All are not equal: An examination of the economic returns to different forms of participation in open source software communities. *Inform. Systems Res.* 24(3):520–538.

Harper FM, Raban D, Rafaeli S, Konstan JA (2008) Predictors of answer quality in online Q&A sites. *Proc. SIGCHI Conf. Human Factors Comput. Systems* (ACM), 865–874.

Holmstrom B, Milgrom P (1991) Multitask principal-agent analyses: Incentive contracts, asset ownership, and job design. *J. Law Econom. Organ.* 7:24–52.

Hosanagar K, Fleder D, Lee D, Buja A (2013) Will the global village fracture into tribes? Recommender systems and their effects on consumer fragmentation. *Management Sci.* 60(4):805–823.

Hsieh G, Kraut RE, Hudson SE (2010) Why pay? Exploring how financial incentives are used for question & answer. *Proc. SIGCHI Conf. Human Factors Comput. Systems* (ACM), 305–314.

Iacus SM, King G, Porro G (2012) Causal inference without balance checking: Coarsened exact matching. *Political Anal.* 20(1):1–24.

Judge TA, Bretz RD Jr (1994) Political influence behavior and career success. *J. Management* 20(1):43–65.

Kachelmeier SJ, Reichert BE, Williamson MG (2008) Measuring and motivating quantity, creativity, or both. *J. Accounting Res.* 46(2):341–373.

Khansa L, Ma X, Liginlal D, Kim SS (2015) Understanding members' active participation in online question-and-answer communities: A theory and empirical analysis. *J. Management Inform. Systems* 32(2):162–203.

Khern-am-nuai W, Kannan KN, Ghasemkhani H (2018) Extrinsic vs. intrinsic rewards for contributing reviews in an online platform. *Inform. Systems Res.* 29(4):871–892.

Kuang L, Huang N, Hong Y, Yan Z (2019) Spillover effects of financial incentives on non-incentivized user engagement: Evidence from an online knowledge exchange platform. *J. Management Inform. Systems* 36(1):289–320.

Lange D, Lee PM, Dai Y (2011) Organizational reputation: A review. *J. Management* 37(1):153–184.

Lerner J, Tirole J (2002) Some simple economics of open source. *J. Indust. Econom.* 50(2):197–234.

Li Z, Huang KW, Cavusoglu H (2012) Quantifying the impact of badges on user engagement in online Q&A communities. *33rd Internat. Conf. Inform. Systems.*

Lin HF (2007) Effects of extrinsic and intrinsic motivation on employee knowledge sharing intentions. *J. Inform. Sci.* 33(2):135–149.

Lou J, Fang Y, Lim KH, Peng JZ (2013) Contributing high quantity and quality knowledge to online Q&A communities. *J. Assoc. Inform. Sci. Tech.* 64(2):356–371.

Mamykina L, Manoim B, Mittal M, Hripcsak G, Hartmann B (2011) Design lessons from the fastest Q&A site in the west. *Proc. SIGCHI Conf. Human Factors Comput. Systems* (ACM, New York), 2857–2866.

Meier S (2007) Do subsidies increase charitable giving in the long run? Matching donations in a field experiment. *J. Eur. Econom. Assoc.* 5(6):1203–1222.

Mochon D, Schwartz J, Maroba J, Patel D, Ariely D (2016) Gain without pain: The extended effects of a behavioral health intervention. *Management Sci.* 63(1):58–72.

Overby E, Forman C (2014) The effect of electronic commerce on geographic purchasing patterns and price dispersion. *Management Sci.* 61(2):431–453.

Pu J, Kwark Y, Han S, Gu B, Ye Q (2017) The double-edged sword of expert reviewer programs: The effects of offering expert reviewer status on review generation. *38th Internat. Conf. Inform. Systems.*

Rui H, Whinston A (2012) Information or attention? An empirical study of user contribution on Twitter. *Inform. Systems E-Bus. Management* 10(3):309–324.

Shadish W, Cook TD, Campbell DT (2002) *Experimental and Quasi-Experimental Designs for Generalized Causal Inference* (Houghton Mifflin, Boston).

Stephen A, Bart Y, Du Plessis C, Goncalves D (2012) *Does Paying for Online Product Reviews Pay Off? The Effects of Monetary Incentives on Content Creators and Consumers* (ACR North American Advances, Duluth, MN).

Sun M, Zhu F (2013) Ad revenue and content commercialization: Evidence from blogs. *Management Sci.* 59(10):2314–2331.

Sun Y, Dong X, McIntyre S (2017) Motivation of user-generated content: Social connectedness moderates the effects of monetary rewards. *Marketing Sci.* 36(3):329–337.

Tang Q, Gu B, Whinston AB (2012) Content contribution for revenue sharing and reputation in social media: A dynamic structural model. *J. Management Inform. Systems* 29(2):41–76.

Titmuss RM (1970) *The Gift Relationship* (Allen and Unwin, London).

Wang J, Ghose A, Ipeirotis P (2012) Bonus, disclosure, and choice: What motivates the creation of high-quality paid reviews? *33rd Internat. Conf. Inform. Systems.*

Wasko MM, Faraj S (2005) Why should I share? Examining knowledge contribution in networks of practice. *Management Inform. Systems Quart.* 29(1):35–57.

Wayne SJ, Ferris GR (1990) Influence tactics, affect, and exchange quality in supervisor-subordinate interactions: A laboratory experiment and field study. *J. Appl. Psych.* 75(5): 487–499.

Weiss A (1995) Human capital vs. signalling explanations of wages. *J. Econom. Perspect.* 9(4):133–154.

Xu L, Nian T, Cabral L (2020) What makes geeks tick? A study of stack overflow careers. *Management Sci.* 66(2):587–604.

Yu J, Jiang Z, Chan HC (2007) Knowledge contribution in problem solving virtual communities: The mediating role of individual motivations. *Proc. 2007 ACM SIGMIS CPR Conf. Comput. Personnel Res. Global Inform. Tech. Workforce* (ACM, New York), 144–152.

Zhang X, Chen W, Gu B, Guo X (2017) The impact of introductory incentives on new users: Evidence from an online health community. *38th Internat. Conf. Inform. Systems*. (Elsevier, Amsterdam).