Ad Revenue and Content Commercialization: Evidence from Blogs*

Monic Sun
Graduate School of Business
Stanford University
Stanford, CA 94305-5015
Phone: +1 (650) 721-1265
Email: monic@stanford.edu

Feng Zhu
Marshall School of Business
University of Southern California
Los Angeles, CA 90089-0808
Phone: +1 (213) 740-8469
Email: fzhu@marshall.usc.edu

March 19, 2011

*We thank Paul Adler, Ron Adner, Victor Bennett, Erik Brynjolfsson, Giovanni Gavetti, Lisa George, Avi Goldfarb, Yinghua He, Jim Lattin, Michael Manove, Kyle Mayer, Nandini Rajagopalan, Marc Rysman, Catherine Tucker, Dennis Yao, participants of the *O&S* workshop and the Applied Economics workshop at the University of Southern California, and seminar participants at Stanford University, the University of British Columbia, UC Davis and the Sixth Bi-annual Conference on the Economics of Intellectual Property, Software and the Internet for helpful discussion and insightful comments. We also thank *eaglet*, the author of the open source software *Pan Gu Segment*, for answering many questions related to the use of the software. Finally, we thank Yinfeng Qin and Rita Ren for research assistance.

Ad Revenue and Content Commercialization: Evidence from Blogs

Abstract

Many scholars are concerned about the impact of ad-sponsored business models on content providers. They argue that content providers, when incentivized by ad revenue, are more likely to tailor their content to attract "eyeballs," and as a result, popular content may be excessively supplied. We empirically test this prediction by taking advantage of the launch of an ad revenue-sharing program initiated by a major Chinese portal site in September 2007. Participating bloggers allow the site to run ads on their blogs and receive 50% of the revenue generated by these ads. After analyzing 4.4 million blog posts, we find that compared to nonparticipants, the percentage of popular content increases by about 12.8 percentage points on the participants' blogs after the program takes effect. This increase can be partly attributed to topics shifting towards three domains: stock market, salacious content, and celebrities. We also find evidence that, relative to nonparticipants, the participants' content quality increases after the program takes effect.

1 Introduction

Many scholars criticize the use of ad-sponsored business models in media industries (e.g., Cross 1994; Herman and McChesney 1997; Turow 1998; McChesney 2004; Anderson and Gabszewicz 2006). They argue that when supported by advertising revenue, media firms have incentives to cater their content production to popular tastes so that they can attract the maximal number of eyeballs. As a result, popular content will be duplicated and excessively supplied, leaving those viewers with niche preferences under-served (e.g., Anderson and Gabszewicz 2006). Anecdotal evidence generally supports these criticisms. Broadcast television networks in the US, for example, are frequently blamed for abolishing advertising-unfriendly programs and sticking with redundant ones (e.g., Goettler and Shachar 2001; Brown and Cavazos 2003; McChesney 2004; Wilbur 2008). Similarly, most newspapers and magazines are charged with being designed for advertising rather than fundamental editorial content (Bagdikian 2004, pp. 241-246).

To make matters worse, popular content is often not the most consequential to readers and may promote unintended social norms. For example, as violence and sex in general sell well, many content providers routinely employ them even though the consequences can be socially detrimental (Herman and McChesney 1997, p. 137). Although regulations such as the *Fairness Doctrine* require commercial broadcasters to present ample issues of public importance, these regulations were never enforced (McChesney 2004, p. 44). Because media are important drivers of culture, some critics went so far as to argue that the advertising-media relationship is effectively destroying the culture and that we in society are "amusing ourselves to death" (e.g., Turow 1998; Postman 2005). The problem is just as severe in developing countries as it is in the developed world (e.g., Zhang 2007). The Chinese speed-dating show, "If You are the One," for example, was the most popular show in the country in 2010.[1] The show was so salacious, materialistic, and popular that the Chinese government

[1] CSM Media Research (`http://www.csm.com.cn`) and `http://news.sina.com.cn/m/news/roll/`

decided to intervene shortly after its launch,[2] and it was suspended and restructured before it was broadcasted again.

Similar to offline media, ad-sponsored business models are pervasive in online media today. Many content sites such as the *New York Times* and Hulu rely entirely on ad revenue to finance their operations. In fact, ad-sponsored business models are no longer limited to corporations. Individual content providers today can also earn ad revenue from the content they provide. The most popular video-sharing site, YouTube, for example, started sharing ad revenue with its top contributors in 2007, and it recently extended the revenue-sharing program to all contributors.[3] Many other content sites based on user-generated content (e.g., blog[4] sites such as Blogger and WordPress) have adopted similar practices.

Given the growing pervasiveness of ad-sponsored business models in media industries today, it is important to examine the extent to which such models shift content providers' incentives. While the theoretical literature in economics repeatedly find support for the claim that ad revenue induces content providers to produce popular or mainstream content (Steiner 1952; Beebe 1977; Spence and Owen 1977; Gal-Or and Dukes 2003; Anderson and Coate 2005; Gabszewicz et al. 2006; Peitz and Valletti 2008), this prediction has received surprisingly little empirical evaluation. The lack of empirical evidence is perhaps because of the difficulty in establishing a causal relationship: If providers of popular content are more likely to seek ad revenue, then the causal relationship could be in the opposite direction.[5]

In this study, we empirically evaluate the impact of ad-sponsored business models on content providers' incentives by taking advantage of the introduction of an ad-revenue-sharing

`2010-10-16/150221288881.shtml`, both accessed December 2010.

[2] Both the show and the intervention gathered tremendous publicity. See an article from the *New York Times* (`http://www.nytimes.com/2010/07/19/world/asia/19chinatv.html`) for more details.

[3] `http://www.youtube.com/partners`, accessed September 2010.

[4] A blog (a blend of the term web log) is a type of website or part of a website. Blogs are usually maintained by an individual with regular entries of commentaries, descriptions of events, or other material, such as pictures or video clips. Entries are commonly displayed in reverse-chronological order.

[5] In a similar vein, Kind et al. (2009) show that the degree of content differentiation between media firms' products may affect their dependency on ad revenue.

program by a major Chinese portal site in September 2007. Participating bloggers allow the site to run ads on their blogs, and in return, they receive 50% of the revenue generated by these ads. We use a differences-in-differences approach to compare content shift of 4,200 participants before and after the program takes effect to that of nonparticipants. We also employ fixed-effects and instrumental-variables approaches to account for bloggers' endogenous decisions to participate in the program. After analyzing 4.4 million blog posts, we find that relative to nonparticipants, the percentage of popular content increases by about 12.8 percentage points on the participants' blogs after the revenue-sharing program takes effect. This increase can be partly attributed to topics shifting towards three domains: stock market, salacious content, and celebrities. At the same time, we find a significant quality improvement of participants' blog posts.

Broadly, this paper also contributes to the growing literature in economics examining factors that influence media content. Scholars have examined how the positioning of media content is affected by the entry of national media (e.g., George and Waldfogel 2006) and the mix of consumer types (e.g., George and Waldfogel 2003), how content quality changes with the emergence of the Internet (e.g., Frijters and Velamuri 2010) and competition intensity (e.g., Zaller 1999; Gentzkow and Shapiro 2008), and how content variety changes as media firms consolidate (e.g., Berry and Waldfogel 2001; George 2002, 2007). They have also identified sources of media bias such as journalists' desire to enhance their career opportunities (Baron 2006) and readers' desire for reinforcement of their prior beliefs (e.g., Groseclose and Milyo 2005; Mullainathan and Shleifer 2005; Gentzkow and Shapiro 2006; Bernhardt et al. 2008; Gal-Or et al. 2010; Gentzkow and Shapiro 2010). Our paper complements these studies by providing empirical evidence on the impact of business models on media content.

The paper proceeds as follows. Section 2 provides details on the empirical setting. Section 3 describes the data. Section 4 presents empirical results and robustness checks, and Section 5 concludes.

2 Background

Our empirical setting is a major Chinese portal site. The site offers many different types of services, including news, emailing, blogging, photo and video-sharing, microblogging, and instant messaging. Our analysis focuses on its blogging businesses. The site started to host blogs for free in September 2005. It is considered a late mover in the blogging business, as the first Chinese blog-hosting site appeared in 2002,[6] and since then many other websites have started providing blogging services. For the first two years, the portal site did not place any ads on individual bloggers' content pages. Then, on September 11, 2007, the company behind the portal site announced the ad-revenue-sharing program. Outsiders were not aware of this program before the announcement date, as the company had kept the development of the program strictly confidential to avoid competitive responses by its rivals.

From September 2007 to March 2008, the company conducted a test run of the ad-revenue-sharing program and invited about 3,000 bloggers to participate. About 1,000 bloggers joined the program during this period. In April 2008, the test period ended and the site started to accept applications from all bloggers. As indicated in the application guidelines, for an application to be successful, the blog needs to have a minimum of 700 visits per week for four consecutive weeks prior to the application date. Once approved, the site places ads on the blog and the blogger receives 50% of the ad revenue generated by the traffic to her blog pages. To participate in the program, the blogger also needs to provide the site with personal information, such as her real name, home address, and bank information. Payments are deposited to participants' bank accounts on a monthly basis whenever the balance exceeds RMB¥100 (equivalent to about US$15).

On the advertiser side, the site uses a pay-per-impression mechanism: At the beginning of each quarter, it announces a fixed price per thousand impressions, and advertisers decide on

[6] http://www.bokee.com, accessed in December 2010.

the amount of impressions to purchase. The site started selling impressions in October 2007, one month after the program's announcement. In November 2007, program participants started to notice ads on their blog pages. Bloggers cannot choose the specific ads to be displayed on their blogs. While the site tries to match the ads with the content of the blog posts, bloggers receive the same amount of money for each impression at a given time. At the beginning of the revenue-sharing program, a blogger would make RMB¥2 (equivalent to about US$0.3) per one thousand impressions generated by her blog posts. To avoid annoying viewers, ads are displayed as a small pop-up window on the lower-right corner of the screen, and the pop-up window automatically disappears within 2 to 3 seconds after the web page finishes loading.

The blog-hosting site offers an ideal setting for our study for multiple reasons. First, the site is one of the largest in China. When the revenue-sharing program was introduced, blogs on this site generated about 0.2 billion page views per day, and on a single day, a popular blog post could receive more than 100,000 page views. Given the amount of attention the blogs receive, any systematic change in the content is economically important. Second, unlike many video-sharing sites, our target site offers unlimited storage space to content providers.[7] As a result, bloggers have little incentive to delete their old posts, allowing us to collect data on the complete history of blog posts from each blogger in our sample. Third, perhaps the most important advantage of this empirical setting is the change in the site's business model: It did not compensate content providers initially, but suddenly introduced the revenue-sharing program. The setting enables us to observe the change in content production for each participant and thus estimate the *influence* of ad-sponsored models on the content providers' incentives. As not every blogger participated in the revenue-sharing program, we can use those nonparticipants as a control group in our analysis. Finally, as the site uses

[7] The only limitation is that a single blog post cannot have more than 20,000 Chinese characters. In Chinese, characters form the basic unit of meaning. Most Chinese words are formed by two or three characters.

the pay-per-impression mechanism on the advertiser side, we do not have to worry about differences among advertisements and the possibility that bloggers tweak their content to target different audiences to get a higher click-through rate.

3 Data

The company that runs the site provided us with a data set that contains a complete list of all bloggers enrolled in the ad-revenue-sharing program as of January 31, 2009 and the dates each blogger joined the program. Each blogger is associated with a unique ten-digit ID. In total, our data include 5,140 participants, among which 4,200 joined the program after April 2008. We focus our analysis on the bloggers who joined after April 2008, as the motivations of invited participants during the test period could be different. Figure 1 shows the number of bloggers enrolled in the program in each month since April 2008. More than 1,700 bloggers enrolled in the program right after it became open to the general public, and after that, a few hundred bloggers enrolled in the program every month.

To control for general trends in the content for all blogs, we create a control group by randomly generating another 50 million ten-digit ID numbers. Many of these IDs are mapped to users without blogs: They are users of the portal site's other services. For the bloggers, we drop those who started blogging after January 2009, and then select the bloggers who write more than one blog post per month, on average. We apply this last criterion to focus our analysis on active bloggers: Many bloggers only create one or two, often very short, posts right after starting blogging and never blog again. It seems that these bloggers want to experience what blogging is like but are not serious about producing any content. In the end, we obtain a list of 26,974 nonparticipants.

We download every blog post that each of the 4,200 participants and 26,974 nonparticipants had written on the site. For each blog post, we collect information on the date it was posted, the title, the number of characters, pictures, and videos in the post, and how many times the post had been read and bookmarked by its viewers. We also collect the tags supplied by the bloggers for each post.

We focus our analysis between May 2007 and January 2009 for two reasons. First, as we rely on tags to identify popular topics for each month in China, it is critical that we aggregate tags across a sufficient number of bloggers in each month in our sample. After its launch in September 2005, the site's blogging service experienced accelerated growth in 2006, and by the first quarter of 2007, it became the largest blog-hosting site (by the number of visitors) in China. In addition, the site did not introduce the tagging feature until April 2007. In May 2007, around 50% blog posts in our sample have tags and this percentage increased to more than 90% in January 2009. For blog posts with no tags, we use post titles to generate tags.[8]

From May 2007 to January 2009, the bloggers in our dataset composed 4,359,197 blog posts. These eventual participants in the program contributed 1,904,609 (43.7%) posts and nonparticipants contributed 2,454,588 (56.3%) posts. Figure 2 shows the average number of blog posts in each month by participants and nonparticipants. We find that participants blogged much more frequently, on average, than nonparticipants. Participants also increased the number of blog posts over time, and the increase was most pronounced when the revenue-sharing program became open to all bloggers. In contrast, nonparticipants' average number of blog posts declined slightly over time. The pattern suggests that the program indeed motivated participants to produce more content. We also find that the average number of blog posts dropped significantly in February 2008 and January 2009. This is most likely

[8]We use *Pau Gu Segment*, an open source software that divides Chinese sentences into a set of keywords, to generate these tags. The software is based on a library of more than 170,000 Chinese keywords and has been used by many commercial firms to build Chinese search engines.

because of the Chinese New Year.[9]

We now consider the popularity of these blog posts. A natural way to consider post popularity is to check whether the post is associated with a popular tag. To gauge interest on the tags, we define a tag's popularity in a certain month by the total number of page views of blog posts containing the tag in that month.[10] On average, we have 59,132 tags per month. We rank all the tags based on their popularity and consider, in each month, the top 150 tags as popular tags. We choose this threshold in order to have a reasonable set of popular tags. Other thresholds such as 100, 500, or top 0.5% or 1% of all tags provide similar results. The tag "stock market" is the most popular tag in most months. We then identify blog posts associated with the popular tags in each month as popular posts. On average, 23% of all the posts in our data are classified as popular posts, and these popular posts on average obtain 56% of the page views in each month.

Next, we compute the percentage of popular blog posts for each blogger in each month. Figure 3 shows how this percentage evolves over time for both participants and nonparticipants in our data. We find that, on average, participants are more likely to post popular content. The percentage of popular content for participants and nonparticipants diverged even more upon the launch of the program: While the percentage of popular content for nonparticipants stayed around 13%, the percentage for participants had a small increase upon the introduction of the program and a significant increase when the program became open to all bloggers. We also notice month-specific effects on the percentage of popular content for both participants and nonparticipants. In May 2008, for example, the percentage of popular content for all bloggers had a sudden increase. This increase resulted from the Wenchuan earthquake, which occurred on May 12, 2008 in China's Sichuan province and killed more

[9]The dates for the Chinese New Year in these two years are February 7, 2008 and January 26, 2009. The 7-10 holidays around the Chinese New Year are typically marked by family gatherings and visits to relatives and friends.

[10]It is important to analyze data on a monthly basis, as a tag's popularity may change over time. The tag "Chinese New Year," for example, is popular only at the beginning of a year.

than 69,000 people. The earthquake was the most discussed topic on TV and in newspapers in that month, and the tag "earthquake" was the most popular one in that month in our data. The percentage dropped back to its average level for the nonparticipating group right after May 2008 but remained at a high level for the participating group. Similarly, the increase in the percentage of popular content for both groups in August 2008 resulted from the opening of the Summer Olympic Games in Beijing.

4 Regression Analysis

We now turn to regression frameworks to detect shifts in the different aspects of blog content for program participants relative to those for nonparticipants.

4.1 Shift in Content Popularity

We first consider content popularity. We employ a differences-in-differences approach with the specification below:

$$\begin{aligned} \%\ Popular_{it} &= \beta_0 + \beta_1 EventualParticipant_i + \beta_2 EventualParticipant_i \times After_{it} \\ &\quad + \sum_{j=2}^{21} \gamma_j MonthDummy_j + \epsilon_{it}, \end{aligned} \tag{1}$$

where $\%\ Popular_{it}$ is the percentage of popular blog posts contributed by blogger i in month t. $EventualParticipant_i$ is a dummy that takes the value of 1 if blogger i is an eventual participant in the program, and 0 otherwise. It captures the systematic difference between program participants and nonparticipants. $After_{it}$ is 1 if blogger i is an eventual participant and has already enrolled in the program in month t, and 0 otherwise. β_2 is our differences-in-differences estimator that captures the effect of the revenue-sharing program on content popularity for participants. We also include dummies for each month, from May 2007 to

January 2009, to control for changes in all bloggers' propensity to produce popular content. May 2007 is used as the benchmark month and is thus dropped from the regression.

We need to address two problems in our specification. First, we need to account for a potential endogeneity problem, as those who apply to join the revenue-sharing program are not randomly selected. In other words, some unobserved heterogeneity among bloggers in the error term may be correlated with their decisions to participate in the program, leading to biased estimates. The participants, for example, could in general like to blog about popular topics more. As a result, their blog posts are more popular and it is easier for them to qualify for the revenue-sharing program.

We take two approaches to address this problem. First, we introduce blogger-level fixed effects to control for time-invariant, unobserved blogger characteristics. Fixed effects allow us to focus on changes in content popularity over time for any given blogger, rather than the absolute levels. As fixed effects do not control for time-variant factors that may be correlated with the decision to participate in the program, we also construct two instrumental variables by taking advantage of the minimum number of page views required to participate in this program. Valid instruments need to correlate with the decision to participate and affect the dependent variable (*% Popular*) only through the participation decision. Our first instrument is the number of months since a blogger's first post. Our second instrument is the average number of posts per month for the blogger in the past. The idea is that the longer a blogger has been blogging or the more frequently she published posts in the past, the more likely that she had cultivated an audience base with more than 700 page views per week, which would make her eligible for the program. At the same time, the two variables are unlikely to be directly correlated with the percentage of popular posts.[11]

Second, our regressions could underestimate the revenue-sharing program's impact. Blog-

[11] One might worry that blogging early on the site may be an indication that the blogger likes popular things. This concern is alleviated by the fact that this site is a late mover in the blogging business and those who like popular things, such as new technologies, are likely to have started blogging on other sites.

gers may tailor their content to improve their chance of getting approved for the revenue-sharing program, and hence the revenue-sharing program would take effect before they join the program. More generally, as it takes time to increase the popularity of one's blog, some bloggers may choose to shift toward popular content right after the program's announcement and wait until their page views meet the requirement before applying to the program. Indeed, the increase in content popularity for the participants right after the program's announcement, as shown in Figure 3, suggests that such effects may have taken place. On top of that, some nonbloggers might be incentivized by the revenue-sharing program to start blogging by focusing on popular content. As a result, bloggers with start dates after September 2007 could be systematically different from those who joined before the program's announcement. Finally, some nonparticipants may also be incentivized by ad revenue. They may have increased their content popularity but still have not met the program requirement.

To minimize these effects, we repeat our analysis after taking September 2007 as the breakpoint for all participants and including only those bloggers who started blogging on the site before September 2007. To ensure that the announcement of the program is truly exogenous, we search baidu.com, the top search engine in China, for news related to the revenue-sharing program. All news is dated on or after the day of the program's announcement. We also search the text of all blog posts in our dataset, as bloggers on the site are likely to discuss this program once they become aware of it. All posts that mention this program are also dated after the program's announcement.

Table 1 summarizes our regression results. In the first three models, we use bloggers' enrollment date as break points. Model (1) reports the results based on ordinary least square (OLS) regression. On average, a participating blogger's percentage of popular posts, before she joins the program, is higher than that of a nonparticipant by 22.0 percentage points. This percentage increases by an additional 7.0 percentage points after she joins the program. Consistent with Figure 3, we also find that bloggers are more likely to produce popular content during May and August 2008 than in other months. Model (2) reports the results with fixed effects, which are similar to those in Model (1). The variable, $EventualParticipant_i$, drops from the regression, as its value does not vary over time. Model (3) reports the results with both fixed effects and instrumental variables. The results in Model (3) show that the revenue-sharing program's effect becomes stronger after correcting for endogeneity, suggesting that, ceteris paribus, if bloggers were randomly chosen to join the program, the shift toward popular content would be even greater than what we observed in the actual data.

In the next three models, we repeat the analysis in Models (1)-(3) using September 2007 as the break point for all participants. We redefine our dummy variable $After_{it}$ to be 1 if month t is on or after September 2007, and 0 otherwise. We find that the systematic difference in the percentage of popular posts between participants and nonparticipants becomes smaller (16.0%). As expected, the impact of the revenue-sharing program is more pronounced: An eventual participant's percentage of popular posts increases by as much as 12.8 percentage points after the program's announcement. The results are consistent with our conjecture that after the program's announcement, many participants started providing popular content in preparation for enrolling in the program.

4.2 Shift in Content Topics

We now examine the shift in topics of participants' blog posts after the program takes effect. After speaking with several frequent bloggers in China, we decide to focus on three most-mentioned topics: stock market, salacious content, and celebrities. China's stock market started in early 1990 and has been notorious for its fluctuations. To maximize their returns, many people regularly read blog posts related to the stock market for free opinions and recommendations. Hence, blogging about the stock market is likely to be an effective strategy in attracting traffic. The other two, salacious content and celebrities, are universally considered as hot topics.[12]

Two research assistants independently examined the top 150 tags in each month and classified each tag into one of four domains: stock market, salacious content, celebrities, and others. The results are highly consistent and the few discrepancies were resolved by a meeting of the research assistants. On average, in each month, 12% of the popular tags (e.g., "stock market index" and "stock recommendation") are classified as being related to the stock market, 13% (e.g., "nude photo scandal" and names of Japanese adult-video idols) are classified as being salacious content, 9% (e.g., "celebrity gossip" and names of the celebrities) are classified as being related to celebrities, and 66% (e.g., "earthquake' and 'Chinese New Year') are classified as others.

We then classify blog posts based on the tags with which they are associated. A blog post may be classified into multiple domains. For example, a post on a nude photo scandal involving celebrities is classified as both salacious and related to celebrities. In the end, 6.2%, 6.5% and 3.0% of the posts are classified as being related to stock market, salacious content and celebrities, respectively. We then compute the percentage of posts in each of these three domains for each blogger in each month, and use these three percentages as dependent variables to repeat the differences-in-differences analysis. Table 2 reports the regression results.

We use September 2007 as the breakpoint for all participants. For each of the three percentages above, we report the results with fixed effects, and the results with both fixed effects and instrumental variables. The results demonstrate significant shifts of content towards all three domains. In total, the blog posts in these three domains for participants increase by 6.7% percentage points (based on the specification with both fixed effects and instrumental variables) relative to nonparticipants.

[12] Although it is illegal in China to post images or videos that contain nudity, or text with explicit descriptions of sexual acts, bloggers can include images or text that are sexually suggestive.

4.3 Shift in Content Quality

Finally, we consider the revenue-sharing program's impact on content quality. As high-quality content attracts eyeballs, participating bloggers may devote more effort to improving the quality of their posts. To identify the extent of such effects, we develop several measures on post quality. For each post, we first compute the percentage of viewers who bookmark the post as one of their favorites.[13] Bloggers cannot tell who bookmarked their posts and hence cannot reciprocate by visiting the blogs of their patrons. Therefore, the only benefit of bookmarking a post is the convenience of re-accessing it in the future. We denote this measure as $\%$ *Bookmark*, which reflects the utility that readers receive from reading the post.

We also measure the amount of each blogger's effort by the average numbers of characters, pictures, and video clips in her posts. Given a blogger, the more effort she devotes into writing, the more likely the blog post has a higher quality. We denote these measures, respectively, as *Num Chars*, *Num Pics*, and *Num Videos*. We take the logarithm of the average number of characters to minimize the effect of outliers. In general, pictures and videos make a post more attractive, although they may also require more effort from the blogger.

We apply the same differences-in-differences approach using each of the four measures above as the dependent variable. Table 3 reports the results. We again use September 2007 as the breakpoint for all participants. We find significant improvement for all measures after the program's announcement. The significant increase in $\%$ *Bookmark* is noteworthy: As bloggers shift toward less differentiated content, readers may find that the blog posts become more repetitive and less worthy of a bookmark. As a result, $\%$ *Bookmark* would have decreased without quality improvement.

[13]For each blog post, the site provides a button such that any reader with an account on the site can bookmark the blog post, which puts the post in her personal collection.

4.4 Additional Robustness Checks

We are concerned about the potential strategic manipulation of tags by program participants. For example, participants may use more tags for their posts after the program's announcement to attract readers. They may also supply popular tags even when these tags are not accurate descriptions of their posts. Such strategic manipulation could have contributed to our finding that program participants are more likely to produce popular content. To address this concern, we generate tags based on the text in each blog post. For each blog post, we tokenize the text into individual words and use the five most frequently mentioned nouns as the tags.[14] We follow the same procedure to identify popular posts and the domains to which each popular post belongs, and repeat the analysis in Tables 1 and 2. The results are similar and in some cases, the effect of the revenue-sharing program is more pronounced.[15]

We are also concerned that some bloggers may write much more frequently than others, and as a result, their blog posts may have a disproportionate influence in determining whether certain content is popular or not: If a small number of prolific bloggers mentions the same tag in every blog post they write, this tag is likely to be classified as a popular tag even if none of the other bloggers use the tag. We repeat our analysis after excluding bloggers whose average monthly number of posts are more than four standard deviations above the mean and obtain similar results.

5 Concluding Remarks

Media consumption today is characterized by two patterns. First, media consumption is moving online, and a significant portion of the content consumed is generated by consumers themselves. All three media sites among the top 10 most-visited websites in the

[14]If a post contains fewer than five nouns, we use the all the nouns as tags.
[15]Regression results for robustness checks are not reported.

world, YouTube, Blogger and Twitter, are based on user contributions[16] (e.g., Dewan and Ramaprasad 2010; George 2008; Gopinath et al. 2010). Second, consumers are increasingly expecting their media consumption to be free (Wray 2010). Given these patterns, content providers are under increasing pressure to monetize their content through ad revenue.

Our study makes the first attempt to empirically evaluate the impact of ad-sponsored business models on user-generated content. We find that, consistent with the theoretical literature, content providers that are sponsored by ad revenue are more likely to generate content that pander to the lowest common denominator. Meanwhile, we also find that ad-sponsored models lead to increased effort on generating content and making it more likeable. The welfare implication of adopting ad-sponsored business model is therefore ambiguous.

In the theoretical literature, quality is often assumed to be exogenous and identical for competing content providers (e.g., Gabszewicz et al. 2006). As a result, these models do not predict the effect of ad-sponsored business models on content quality. Our results suggest that both content quality and location choices need to be endogenized to fully understand the impact of ad-sponsored business models.

Two limitations of this study are noteworthy. First, content providers do not face a capacity constraint in our empirical setting but in offline media, such as radio, television, and newspapers, the capacity constraint can be critical. In such cases, the substitution of less popular content with popular content may be greater when content providers rely more on ad revenue. Second, in China, pornography and politically controversial issues are generally not allowed. Therefore, some of our results could be different in other cultural settings.

[16] Source: `www.alexa.com`, accessed September 2010.

References

Anderson, Simon P., Stephen Coate. 2005. Market provision of broadcasting: A welfare analysis. *Review of Economic Studies* **72**(4) 947–972.

Anderson, Simon P., Jean J. Gabszewicz. 2006. The media and advertising: A tale of two-sided markets. Victor A. Ginsburgh, David Throsby, eds., *Handbook of the Economics of Art and Culture*, vol. 1, chap. 18. Elsevier B.V., 568–614.

Bagdikian, Ben H. 2004. *The New Media Monopoly*. Beacon Press.

Baron, David. 2006. Persistent media bias. *Journal of Public Economics* **90**(1-2) 1–36.

Beebe, Jack H. 1977. Institutional structure and program choices in television markets. *Quarterly Journal of Economics* **91**(1) 15–37.

Bernhardt, Dan, Stefan Krasa, Mattias Polborn. 2008. Political polarization and the electoral effects of media bias. *Journal of Public Economics* **92**(5-6) 1092–1104.

Berry, Steven T., Joel Waldfogel. 2001. Do mergers increase product variety? Evidence from radio broadcasting. *Quarterly Journal of Economics* **116**(3) 1009–1025.

Brown, Keith S., Roberto J. Cavazos. 2003. Empirical aspects of advertiser preferences and program content of network television. Federal Communications Commission Media Bureau Staff Research Paper 2003-1.

Cross, Donna W. 1994. *Media Speak: How Television Makes up Your Mind*. Penguin Group.

Dewan, Sanjeev, Jui Ramaprasad. 2010. Consumer blogging and music sampling: Long tail effects. Working paper, UC Irvine Paul Merage School of Business.

Frijters, Paul, Malathi Velamuri. 2010. Is the Internet bad news? The online news era and the market for high-quality news. *Review of Network Economics* **9**(2).

Gabszewicz, Jean J., Didier Laussel, Nathalie Sonnac. 2006. Competition in the media and advertising markets. *Manchester School* **74**(1) 1–22.

Gal-Or, Esther, Anthony J. Dukes. 2003. Minimum differentiation in commercial media markets. *Journal of Economics & Management Strategy* **12**(3) 291–325.

Gal-Or, Esther, Tansev Grelani, Tuba Pinar Yildirim. 2010. The impact of advertising on media bias. Working paper.

Gentzkow, Matthew, Jesse M. Shapiro. 2006. Media bias and reputation. *Journal of Political Economy* **114**(2) 280–316.

Gentzkow, Matthew, Jesse M. Shapiro. 2008. Competition and truth in the market for news. *Journal of Economic Perspectives* **22**(2) 133–154.

Gentzkow, Matthew, Jesse M. Shapiro. 2010. What drives media slant? Evidence from U.S. daily newspapers. *Econometrica* **78**(1) 35–71.

George, Lisa M. 2002. Ownership concentration and product variety in daily newspaper markets. Lorrie F. Cranor, Shane Greenstein, eds., *Communications Policy and Information Technology: Promises, Problems, Prospects*, chap. 9. MIT Press, Cambridge, MA, 235–253.

George, Lisa M. 2007. What's fit to print: The effect of ownership concentration on product variety in daily newspaper markets. *Information Economics and Policy* **19**(3-4) 285–303.

George, Lisa M. 2008. The Internet and the market for daily newspapers. *The B.E. Journal of Economic Analysis & Policy* **8**(1).

George, Lisa M., Joel Waldfogel. 2003. Who affects whom in daily newspaper markets? *Journal of Political Economy* **111**(4) 765–785.

George, Lisa M., Joel Waldfogel. 2006. The "New York Time" and the market for local newspapers. *American Economic Review* **96**(1) 435–447.

Goettler, Ronald L., Ron Shachar. 2001. Spatial competition in the network television industry. *Rand Journal of Economics* **32**(4) 624–656.

Gopinath, Shyam, Pradeep K. Chintagunta, Sriram Venkataraman. 2010. Measuring local market-level differences in the effects of pre-release blogs on a movie's opening box-office performance. Working paper.

Groseclose, Tim, Jeffrey Milyo. 2005. A measure of media bias. *Quarterly Journal of Economics* **120**(4) 1191–1237.

Herman, Ed S., Robert W. McChesney. 1997. *The Global Media: The Missionaries of Global Capitalism*. Cassell.

Kind, Hans Jarle, Tore Nilssen, Lars Sørgard. 2009. Business models for media firms: Does competition matter for how they raise revenue? *Marketing Science* **28**(6) 1112–1128.

McChesney, Robert W. 2004. *The Problem of the Media: U.S. Communication Politics in the Twenty-First Century*. Monthly Review Press.

Mullainathan, Sendhil, Andrei Shleifer. 2005. The market for news. *American Economic Review* **95**(4) 1031–1053.

Peitz, Martin, Tommaso M. Valletti. 2008. Content and advertising in the media: Pay-tv versus free-to-air. *International Journal of Industrial Organization* **26**(4) 949–965.

Postman, Neil. 2005. *Amusing Ourselves to Death: Public Discourse in the Age of Show Business*. Penguin.

Spence, Michael, Bruce Owen. 1977. Television programming, monopolistic competition, and welfare. *Quarterly Journal of Economics* **91**(1) 103–126.

Steiner, Peter O. 1952. Program patterns and preferences, and the workability of competition in radio broadcasting. *Quarterly Journal of Economics* **66**(2) 194–223.

Turow, Joseph. 1998. *Breaking Up America: Advertisers and the New Media World*. University Of Chicago Press.

Wilbur, Kenneth C. 2008. A two-sided, empirical model of television advertising and viewing markets. *Marketing Science* **27**(3) 356–378.

Wray, Richard. 2010. Media consumption on the increase. The Guardian. April 19, available at http://www.guardian.co.uk/business/2010/apr/19/media-consumption-survey, accessed December 2010.

Zaller, John. 1999. Market competition and news quality. Paper presented at the 1999 Annual Meetings of the American Political Science Association, Atlanta, GA.

Zhang, Bo. 2007. On the vulgarization of mass media. *Press Circles* **4** 66–68.

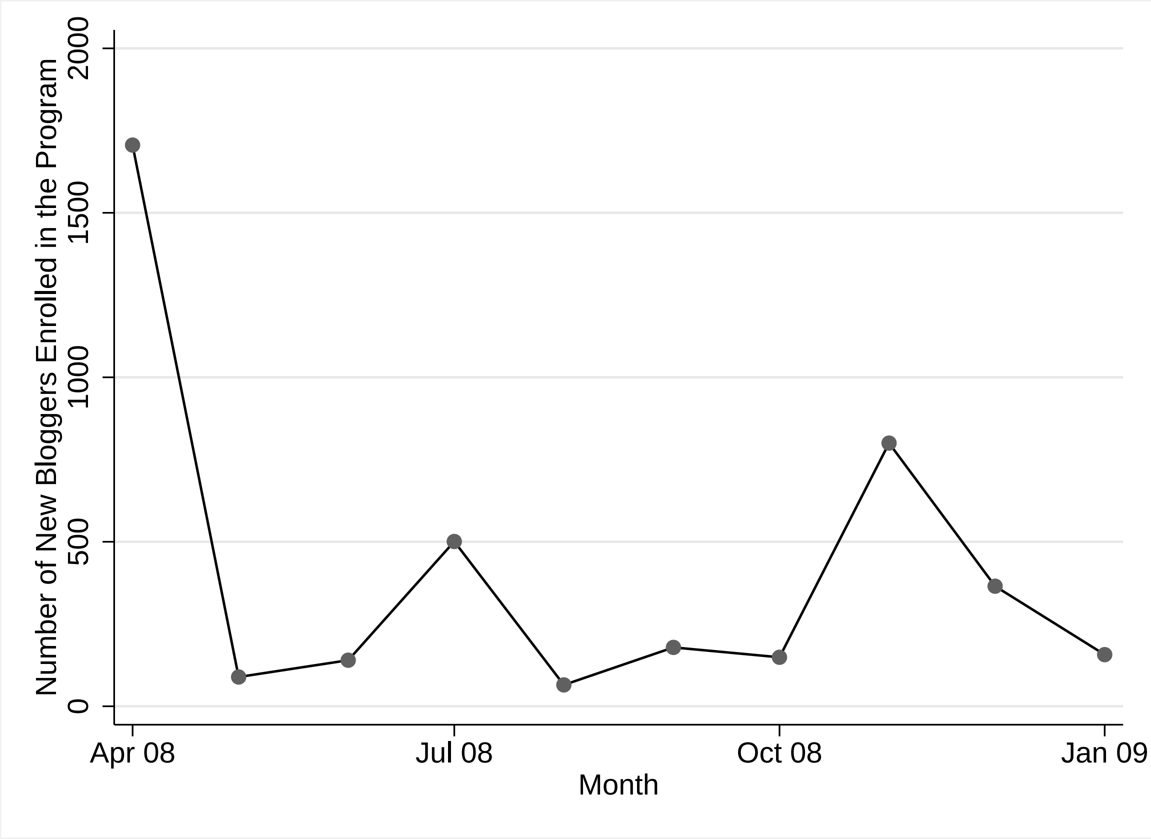
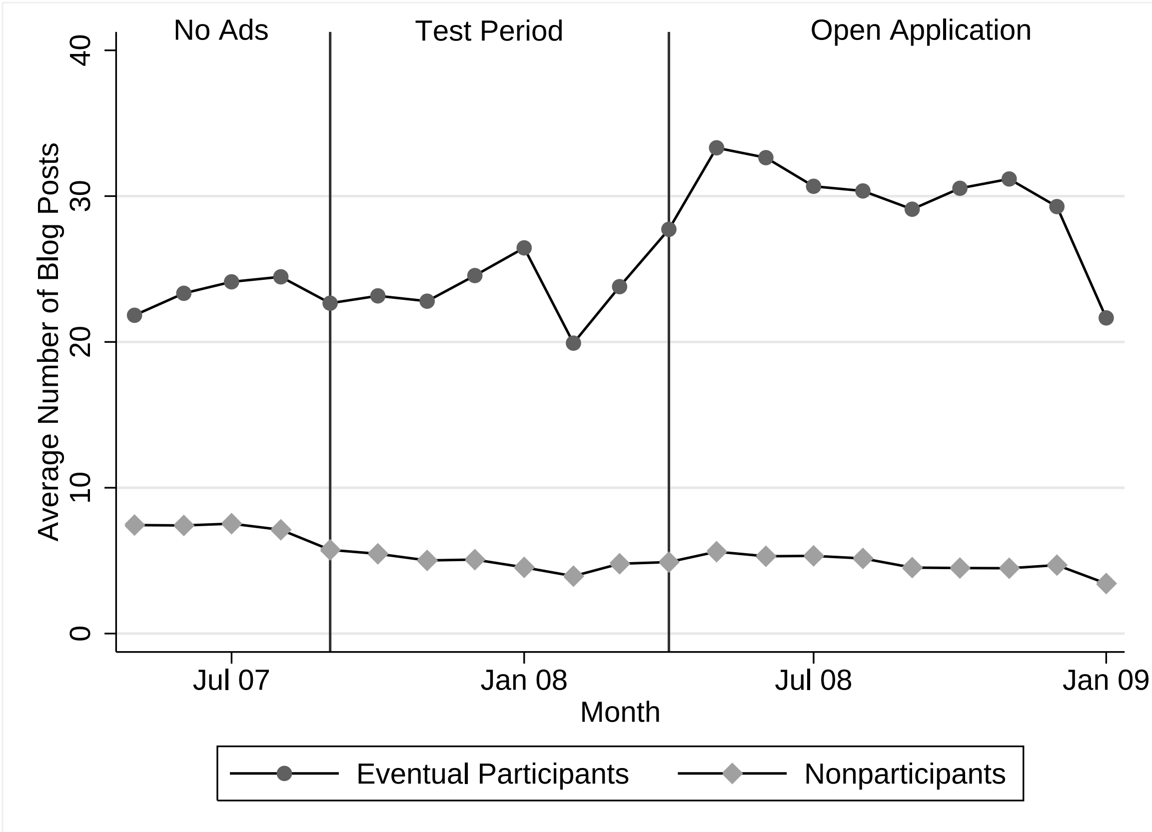
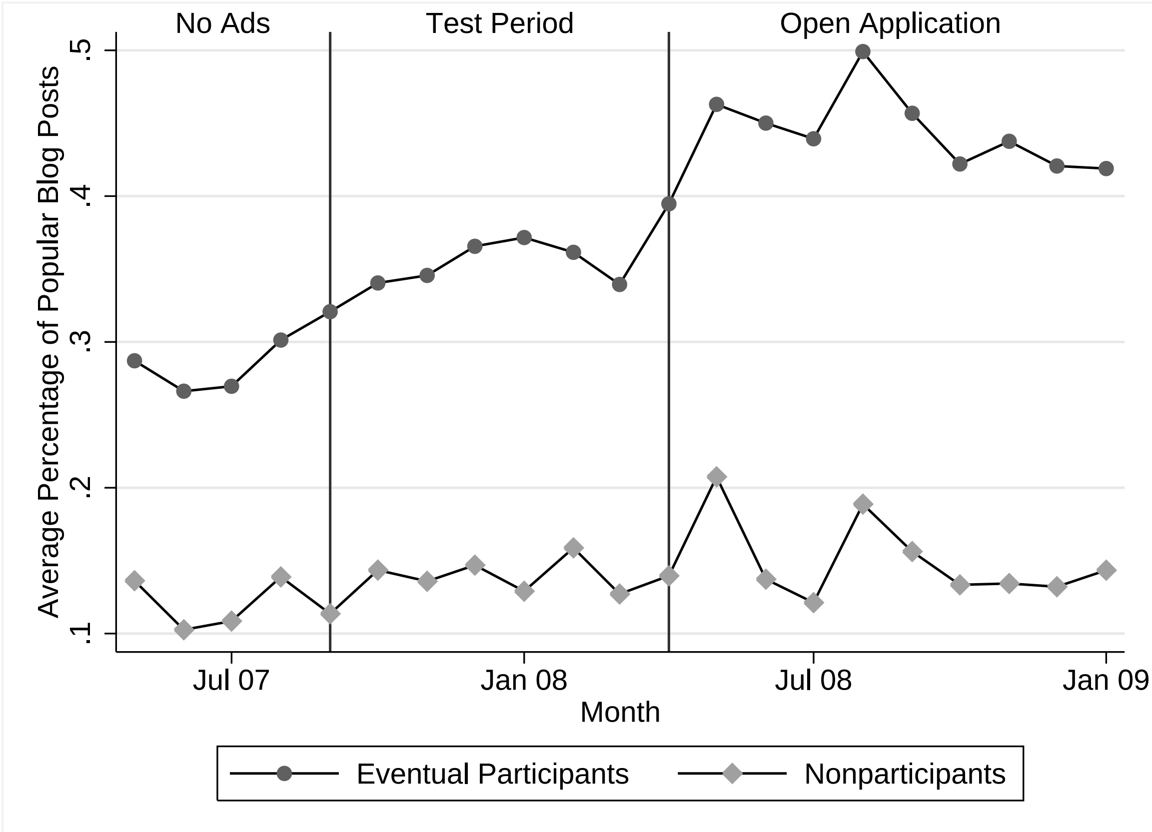


Figure 1: Number of Bloggers Enrolled in the Program in Each Month



Note: The two vertical lines indicate September 2007 and April 2008, respectively.

Figure 2: Average Number of Blog Posts in Each Month



Note: The two vertical lines indicate September 2007 and April 2008, respectively.

Figure 3: Average Percentage of Blog Posts on Popular Topics in Each Month

Table 1: The Impact of Revenue Sharing on Content Popularity

	Enrollment dates as Break Points			9/2007 as the Break Point		
Model	(1)	(2)	(3)	(4)	(5)	(6)
Dependent Variable	% Popular	% Popular	% Popular	% Popular	% Popular	% Popular
EventualParticipant	0.220***			0.160***		
	(0.004)			(0.005)		
EventualParticipant	0.070***	0.070***	0.078***	0.094***	0.091***	0.128***
× AfterEnrollment	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.017)
Dummy for 06/2007	−0.031***	−0.031***	−0.031***	−0.031***	−0.031***	−0.031***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Dummy for 07/2007	−0.026***	−0.025***	−0.025***	−0.026***	−0.026***	−0.026***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Dummy for 08/2007	0.005**	0.005**	0.005**	0.004*	0.004*	0.004*
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Dummy for 09/2007	−0.014***	−0.016***	−0.016***	−0.029***	−0.029***	−0.034***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)
Dummy for 10/2007	0.014***	0.011***	0.011***	−0.003	−0.003	−0.008**
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)
Dummy for 11/2007	0.008***	0.003	0.003	−0.012***	−0.011***	−0.016***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)
Dummy for 12/2007	0.020***	0.013***	0.013***	−0.002	−0.001	−0.007**
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)
Dummy for 01/2008	−0.033***	−0.042***	−0.042***	−0.055***	−0.055***	−0.060***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)
Dummy for 02/2008	0.030***	0.020***	0.020***	0.008***	0.008***	0.003
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)
Dummy for 03/2008	−0.000	−0.012***	−0.012***	−0.026***	−0.026***	−0.031***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)
Dummy for 04/2008	−0.029***	−0.043***	−0.044***	−0.050***	−0.049***	−0.055***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)
Dummy for 05/2008	0.081***	0.065***	0.065***	0.056***	0.057***	0.052***
	(0.002)	(0.002)	(0.002)	(0.003)	(0.002)	(0.003)
Dummy for 06/2008	0.018***	−0.000	−0.001	−0.008***	−0.008***	−0.013***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)
Dummy for 07/2008	0.002	−0.019***	−0.020***	−0.024***	−0.024***	−0.029***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)
Dummy for 08/2008	0.068***	0.046***	0.045***	0.040***	0.041***	0.036***
	(0.002)	(0.002)	(0.002)	(0.003)	(0.003)	(0.003)
Dummy for 09/2008	0.034***	0.010***	0.009***	0.002	0.003	−0.003
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)
Dummy for 10/2008	0.010***	−0.016***	−0.017***	−0.023***	−0.022***	−0.027***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)
Dummy for 11/2008	0.011***	−0.016***	−0.017***	−0.021***	−0.021***	−0.026***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)
Dummy for 12/2008	0.006**	−0.022***	−0.023***	−0.024***	−0.024***	−0.029***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)
Dummy for 01/2009	0.015***	−0.013***	−0.014***	−0.009***	−0.009***	−0.014***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)
Observations	543,796	543,796	543,706	448,560	448,560	448,560
Adjusted R-squared	0.097	0.016	0.016	0.093	0.016	0.016
Number of IDs		31,174	31,084		21,792	21,792
Specification	OLS	FE	FE/2SLS	OLS	FE	FE/2SLS

Note: Heteroskedasticity-adjusted standard errors in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 2: The Impact of Revenue Sharing on Content Topics

Model	(1)	(2)	(3)	(4)	(5)	(6)
Dependent Variable	Stock Market	Stock Market	Salacious Content	Salacious Content	Celebrities	Celebrities
EventualParticipant	0.012***	0.024**	0.024***	0.027***	0.014***	0.016*
× AfterEnrollment	(0.002)	(0.012)	(0.002)	(0.007)	(0.002)	(0.009)
Dummy for 06/2007	0.001	0.001	−0.001	−0.001	0.001	0.001
	(0.000)	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)
Dummy for 07/2007	−0.000	−0.000	−0.001	−0.001	0.002***	0.002***
	(0.000)	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)
Dummy for 08/2007	−0.000	−0.000	0.001	0.001	−0.000	−0.000
	(0.000)	(0.000)	(0.001)	(0.001)	(0.001)	(0.001)
Dummy for 09/2007	−0.002***	−0.004**	−0.006***	−0.007***	−0.003***	−0.004**
	(0.000)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)
Dummy for 10/2007	−0.002***	−0.004**	−0.006***	−0.006***	−0.002***	−0.003*
	(0.000)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)
Dummy for 11/2007	−0.002***	−0.004**	−0.005***	−0.005***	−0.002**	−0.002
	(0.000)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)
Dummy for 12/2007	−0.002***	−0.004**	−0.007***	−0.007***	−0.001*	−0.002
	(0.000)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)
Dummy for 01/2008	−0.001***	−0.003*	−0.005***	−0.006***	−0.002***	−0.002
	(0.000)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)
Dummy for 02/2008	−0.002***	−0.004**	0.009***	0.009***	0.003***	0.003*
	(0.000)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)
Dummy for 03/2008	−0.001***	−0.003*	−0.001	−0.002	−0.001	−0.001
	(0.000)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)
Dummy for 04/2008	−0.001	−0.002	−0.005***	−0.006***	−0.003***	−0.003**
	(0.000)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)
Dummy for 05/2008	−0.002***	−0.004**	0.003***	0.002*	−0.004***	−0.005***
	(0.000)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)
Dummy for 06/2008	−0.000	−0.002	0.002*	0.001	−0.003***	−0.003**
	(0.000)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)
Dummy for 07/2008	−0.001**	−0.003*	0.001	0.000	−0.001	−0.001
	(0.000)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)
Dummy for 08/2008	−0.002***	−0.003*	−0.004***	−0.004***	−0.001	−0.001
	(0.000)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)
Dummy for 09/2008	−0.001	−0.002	−0.007***	−0.007***	−0.002**	−0.002
	(0.000)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)
Dummy for 10/2008	0.003***	0.002	−0.005***	−0.005***	0.003***	0.003*
	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)
Dummy for 11/2008	0.004***	0.002	−0.005***	−0.006***	−0.001	−0.001
	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)
Dummy for 12/2008	0.003***	0.001	−0.005***	−0.006***	0.000	0.000
	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)
Dummy for 01/2009	−0.001	−0.002	−0.011***	−0.011***	−0.003***	−0.003**
	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)
Observations	448,560	448,560	448,560	448,560	448,560	448,560
Adjusted R-squared	0.002	0.001	0.003	0.003	0.001	0.001
Number of IDs	21,792	21,792	21,792	21,792	21,792	21,792
Specification	FE	FE/2SLS	FE	FE/2SLS	FE	FE/2SLS

Note: Heteroskedasticity-adjusted standard errors in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 3: The Impact of Revenue Sharing on Content Quality

Model	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dependent Variable	% Bookmarks	% Bookmarks	Num Chars	Num Chars	Num Pics	Num Pics	Num Videos	Num Videos
EventualParticipant	0.055***	0.096***	1.134***	1.542***	1.490***	2.096***	0.001***	0.001*
× AfterEnrollment	(0.003)	(0.010)	(0.037)	(0.150)	(0.040)	(0.138)	(0.000)	(0.001)
Dummy for 06/2007	0.000	0.000	0.085***	0.083***	−0.064***	−0.068***	−0.000	−0.000
	(0.000)	(0.000)	(0.019)	(0.019)	(0.015)	(0.015)	(0.000)	(0.000)
Dummy for 07/2007	0.001	0.000	0.095***	0.091***	−0.017	−0.023	0.000	0.000
	(0.000)	(0.000)	(0.021)	(0.021)	(0.016)	(0.016)	(0.000)	(0.000)
Dummy for 08/2007	0.001	−0.000	−0.026	−0.032	0.007	−0.002	0.000	0.000
	(0.001)	(0.001)	(0.022)	(0.022)	(0.016)	(0.017)	(0.000)	(0.000)
Dummy for 09/2007	−0.005***	−0.011***	−0.409***	−0.466***	−0.258***	−0.342***	−0.000**	−0.000
	(0.001)	(0.002)	(0.023)	(0.031)	(0.017)	(0.026)	(0.000)	(0.000)
Dummy for 10/2007	−0.006***	−0.011***	−0.504***	−0.561***	−0.190***	−0.275***	−0.000*	−0.000
	(0.001)	(0.002)	(0.024)	(0.032)	(0.019)	(0.027)	(0.000)	(0.000)
Dummy for 11/2007	−0.007***	−0.013***	−0.688***	−0.744***	−0.296***	−0.380***	−0.000	−0.000
	(0.000)	(0.001)	(0.024)	(0.032)	(0.020)	(0.027)	(0.000)	(0.000)
Dummy for 12/2007	−0.007***	−0.012***	−0.740***	−0.797***	−0.362***	−0.446***	−0.000*	−0.000
	(0.000)	(0.001)	(0.025)	(0.033)	(0.018)	(0.026)	(0.000)	(0.000)
Dummy for 01/2008	−0.007***	−0.012***	−0.877***	−0.934***	−0.395***	−0.479***	−0.000*	−0.000
	(0.000)	(0.001)	(0.025)	(0.033)	(0.018)	(0.026)	(0.000)	(0.000)
Dummy for 02/2008	−0.006***	−0.012***	−1.151***	−1.208***	−0.401***	−0.485***	−0.000	−0.000
	(0.001)	(0.002)	(0.026)	(0.034)	(0.019)	(0.026)	(0.000)	(0.000)
Dummy for 03/2008	−0.005***	−0.011***	−0.998***	−1.055***	−0.424***	−0.509***	−0.000*	−0.000
	(0.001)	(0.001)	(0.026)	(0.034)	(0.018)	(0.026)	(0.000)	(0.000)
Dummy for 04/2008	−0.001	−0.007***	−0.970***	−1.027***	−0.263***	−0.347***	0.001***	0.001***
	(0.001)	(0.002)	(0.026)	(0.034)	(0.018)	(0.026)	(0.000)	(0.000)
Dummy for 05/2008	0.001	−0.005***	−0.768***	−0.825***	−0.110***	−0.194***	0.003***	0.003***
	(0.001)	(0.002)	(0.026)	(0.034)	(0.019)	(0.026)	(0.000)	(0.000)
Dummy for 06/2008	0.004***	−0.002	−0.883***	−0.940***	−0.075***	−0.159***	0.002***	0.002***
	(0.001)	(0.001)	(0.026)	(0.034)	(0.019)	(0.026)	(0.000)	(0.000)
Dummy for 07/2008	0.007***	0.001	−0.968***	−1.025***	−0.076***	−0.160***	0.002***	0.002***
	(0.001)	(0.002)	(0.027)	(0.034)	(0.019)	(0.025)	(0.000)	(0.000)
Dummy for 08/2008	0.009***	0.004	−1.148***	−1.205***	0.024	−0.061**	0.002***	0.002***
	(0.002)	(0.002)	(0.027)	(0.034)	(0.021)	(0.027)	(0.000)	(0.000)
Dummy for 09/2008	0.008***	0.002	−1.262***	−1.319***	−0.099***	−0.184***	0.002***	0.002***
	(0.001)	(0.002)	(0.027)	(0.035)	(0.020)	(0.027)	(0.000)	(0.000)
Dummy for 10/2008	0.007***	0.001	−1.317***	−1.374***	−0.018	−0.102***	0.002***	0.002***
	(0.001)	(0.001)	(0.027)	(0.035)	(0.021)	(0.027)	(0.000)	(0.000)
Dummy for 11/2008	0.007***	0.001	−1.483***	−1.540***	−0.150***	−0.235***	0.002***	0.002***
	(0.001)	(0.002)	(0.027)	(0.035)	(0.020)	(0.026)	(0.000)	(0.000)
Dummy for 12/2008	0.006***	0.000	−1.549***	−1.606***	−0.195***	−0.280***	0.002***	0.002***
	(0.001)	(0.002)	(0.028)	(0.035)	(0.019)	(0.026)	(0.000)	(0.000)
Dummy for 01/2009	0.005***	−0.000	−1.833***	−1.890***	−0.272***	−0.357***	0.002***	0.002***
	(0.002)	(0.002)	(0.028)	(0.035)	(0.019)	(0.025)	(0.000)	(0.000)
Observations	448,560	448,560	448,560	448,560	448,560	448,560	448,560	448,560
Adjusted R-squared	0.006	0.004	0.045	0.044	0.015	0.013	0.003	0.003
Number of IDs	21,792	21,792	21,792	21,792	21,792	21,792	21,792	21,792
Specification	FE	FE/2SLS	FE	FE/2SLS	FE	FE/2SLS	FE	FE/2SLS

Note: In Models 1 and 2, we multiple the dependent variable by 100 for the ease of displaying coefficients. Heteroskedasticity-adjusted standard errors in parentheses. * significant at 10%; ** significant at 5%; *** significant at 1%.