American Economic Review 101 (June 2011): 1601–1615
http://www.aeaweb.org/articles.php?doi=10.1257/aer.101.4.1601

Group Size and Incentives to Contribute: A Natural Experiment at Chinese Wikipedia

By Xiaoquan (Michael) Zhang and Feng Zhu*

Many public goods on the Internet today rely entirely on free user contributions. Popular examples include open source software development communities (e.g., Linux, Apache), open content production (e.g., Wikipedia, OpenCourseWare), and content sharing networks (e.g., Flickr, YouTube). Several studies have examined the incentives that motivate these free contributors (e.g., Josh Lerner and Jean Tirole 2002; Karim R. Lakhani and Eric von Hippel 2003).

In this paper, we examine the causal relationship between group size and incentives to contribute in the setting of Chinese Wikipedia, the Chinese language version of an online encyclopedia that relies entirely on voluntary contributions. The group at Chinese Wikipedia is composed of Chinese-speaking people in mainland China, Taiwan, Hong Kong, Singapore, and other regions in the world, who are aware of Chinese Wikipedia and have access to it. Our identification hinges on the exogenous reduction in group size at Chinese Wikipedia as a result of the block of Chinese Wikipedia in mainland China in October 2005. During the block, mainland Chinese could not use or contribute to Chinese Wikipedia, although contributors outside mainland China could continue to do so. We find that contribution levels of these nonblocked contributors decrease by 42.8 percent on average as a result of the block. We attribute the cause to social effects: contributors receive social benefits from their contributions, and the shrinking group size reduces these social benefits. Consistent with our explanation, we find that the more contributors value social benefits, the greater the reduction in their contributions after the block.

The relationship between group size and incentives to contribute is a fundamental question in the literature on the private provision of public goods (e.g., John Chamberlin 1974; Theodore Bergstrom, Lawrence Blume, and Hal R. Varian 1986; James Andreoni 1988a). A major focus of the literature is the free-rider hypothesis: As group size grows, individual contribution levels decline (e.g., Mancur Olson 1965). Theoretical models based on pure altruism generally support the hypothesis (Andreoni 1988a; Timothy L. Fries, Edward Golding, and Richard E. Romano 1991;

* Zhang: Hong Kong University of Science and Technology, HKUST Business School, Clearwater Bay, Kowloon, Hong Kong (e-mail: zhang@ust.hk); Zhu: Marshall School of Business, University of Southern California, Bridge Hall 306, 3670 Trousdale Parkway, Los Angeles, CA 90089-0808 (e-mail: fzhu@marshall.usc.edu). We thank three anonymous referees, Paul Adler, Jim Andreoni, Erik Brynjolfsson, Chris Dellarocas, Philippe De Donder, Shantanu Dutta, Benjamin Edelman, Shane Greenstein, Josh Lerner, Ivan Png, Monic Sun, Qiong Yang and seminar participants at Chinese University of Hong Kong, City University of Hong Kong, Harvard University, Peking University, Tsinghua University, the 2007 Workshop on Information Systems and Economics, the 2010 Wharton Technology Conference, and the Fifth Bi-annual Conference on the Economics of the Software and Internet Industries at the Toulouse School of Economics for valuable comments and suggestions. We also thank "Shizhao," one of the administrators at Chinese Wikipedia, for answering many questions. We gratefully acknowledge the financial support from the NET Institute (www.netinst.org). Both authors contributed equally to this work.

Thomas R. Palfrey and Howard Rosenthal 1984). In these models, a contributor receives utility from the total provision of the public good and her private consumption. As group size grows, the average contribution level falls to zero, and only those individuals with the lowest costs of contributing or the highest income will contribute (Andreoni 1988a; Christopher Bliss and Barry Nalebuff 1984; Marc Bilodeau and Al Slivinski 1996). As Andreoni (1988a) points out, these models fail to explain extensive giving in the charitable sector of the economy.

More recent studies propose models with impure altruism to address the inconsistency between theory and empirical observations (Richard Cornes and Todd Sandler 1984; Richard S. Steinberg 1987; Andreoni 1989, 1990). In these models, a contributor receives not only utility from total provision of the public good, but also a rather selfish, private benefit or warm glow, such as moral satisfaction and joy of giving. The group-size effect depends on how warm glow is introduced into individual contributors' utility function. For example, Andreoni (2007) models a situation in which an individual's utility of giving depends on the number of recipients and finds that the individual's total giving has an ambiguous relationship with group size. More generally, when group size is sufficiently large, the relative importance of pure altruism, which gives rise to free-riding behavior, vanishes and private benefit becomes the dominant motive for contributing (David C. Ribar and Mark O. Wilhelm 2002; Andreoni 2006, p. 1223). Therefore, when private benefit increases with group size, giving rise to "social effects," individual contribution levels could increase with group size in a large group.

Many studies support the idea that contributors to public goods receive private benefits as a result of social effects. For example, scholars show that a contributor's utility of giving may depend on the number of receipts (Andreoni 2007), the amount of giving by other contributors in the social reference group (Andreoni and John Karl Scholz 1998), reciprocity among peers (Andreoni 1988b), the impact of the gift on social welfare (Brian Duncan 2004), prestige (William T. Harbaugh 1998), and the social image they project (Andreoni and B. Douglas Bernheim 2009; Tore Ellingsen and Magnus Johannesson 2008; Adriaan R. Soetevant 2005; Mari Rege and Kjetil Telle 2004; Lakhani and von Hippel 2003; Yuqing Ren, Robert Kraut, and Sara Kiesler 2007). These studies in general find that social effects encourage contribution. They (except Andreoni 2007), however, are conducted in environments with fixed group sizes and thus do not examine the group-size effect.

Empirical studies examining the effect of group size on individual-level contributions are mostly based on experimental data. The group-size effect found in these experiments is ambiguous (John O. Ledyard 1995). John W. Sweeney (1973) and Chamberlin (1978) find increased free-riding behavior in larger groups, while Gerald Marwell and Ruth E. Ames (1979) find the group-size effect to be weak, and Jacob K. Goeree, Charles A. Holt, and Susan K. Laury (2002) find no clear effect. R. Mark Isaac and James M. Walker (1988) and Isaac, Walker, and Arlington W. Williams (1994) examine group sizes of 4, 10, 40, and 100, with varying marginal returns to giving. They find that contributions increase with group size in relatively small groups (from 4 to 10) when the marginal per capita return is high, but the group-size effect vanishes in larger groups with 40 to 100 members. Andreoni (2007) finds that doubling the number of recipients increases but does not double the value of the gift to the giver. Laboratory studies on this topic are necessarily at a

disadvantage, because compared with many groups in the real world, those used in the experiments are very small.

We contribute to the literature by examining the causal relationship between group size and individual-level contribution using field data. Our study provides empirical evidence that social effects indeed dominate free-riding incentives in a setting with a large group size, and thus highlights the importance of social effects in public good provisions.

The rest of the paper is organized as follows. Section I provides a background of Wikipedia and its blocks and unblocks in mainland China. Section II presents the empirical results. Section III conducts various robustness checks. Section IV concludes.

I. Background

Wikipedia is a web-based free encyclopedia project launched in January 2001 and operated by the Wikimedia Foundation, a nonprofit charitable organization.[1] Its goal is to "give every single person in the world free access to the sum of all human knowledge."[2] Wikipedia articles are collectively written by volunteers and can be edited by anyone with access to the Internet. The online encyclopedia contains more than 10 million articles in 264 languages and is the fourth-most-visited website in the world.

Chinese Wikipedia, the Chinese-language edition of Wikipedia, started in October 2002. As a result of its political concerns, the Chinese government implemented the Great Firewall System to censor Chinese mainlanders' access to various information sources, among which are all Wikipedia sites.[3] Access to Chinese Wikipedia has been blocked and unblocked in mainland China five times.[4] Figure 1 depicts the time line of the five blocks on Chinese Wikipedia.

The first block took place on June 2, 2004, when all Wikipedia sites were blocked in mainland China. Two administrators of Chinese Wikipedia, "Shizhao" and "Mountain," contacted their respective Internet service providers (ISPs) and submitted an appeal on June 15, 2004. All Wikipedia sites were unblocked within a week.

The second block, which took place between September 23 and September 27, 2004, was not universal. Only a small number of users in mainland China had problems with accessing Wikipedia.

The third block started on October 19, 2005. "Shizhao" submitted an appeal to his ISP on October 21. Given the experience with the first two blocks, many people expected the block to be lifted soon. This time, the appeal received no response. On the morning of October 31, 2005, surprisingly, contributors from mainland China began to report that they could access Wikipedia. It turned out that this "unblocking" was linked to a server upgrade in the Korean server cluster. A change of the Internet

[1] See Shane Greenstein and Michelle Devereux (2009) for a detailed description of Wikipedia.

[2] http://en.wikipedia.org/wiki/User:Jimbo_Wales (accessed May 2008).

[3] Ronald J. Deibert (2002) and Jonathan Zittrain and Benjamin Edelman (2003) provide a detailed introduction to this issue. Hong Kong and Macau, following the "one country two systems" principle, are not affected by the blocks.

[4] The latest information about the blocks can be found at http://en.wikipedia.org/wiki/Blocking_of_Wikipedia_in_mainland_China, accessed January 2009.

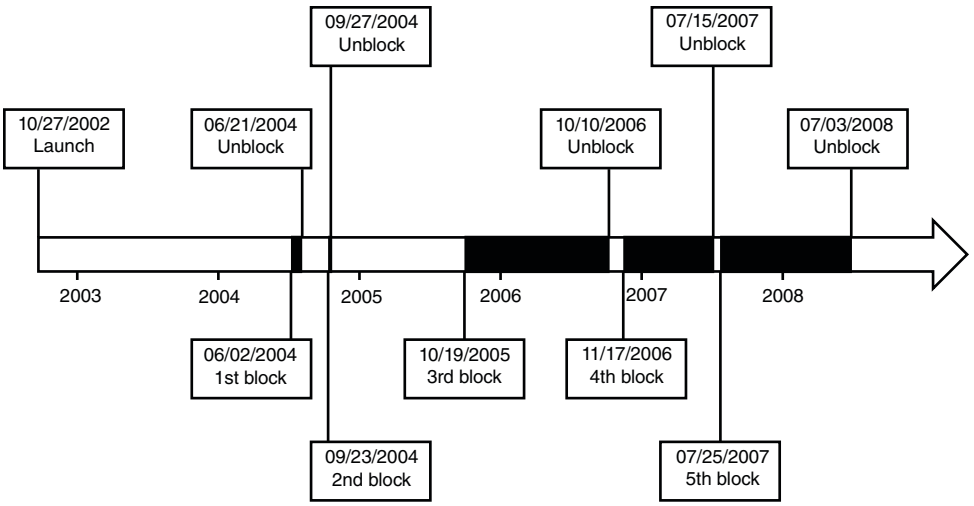


FIGURE 1. TIMELINE OF CHINESE WIKIPEDIA

Notes: The time periods during which Chinese Wikipedia is blocked are indicated in black.

Protocol (IP) address of the Wikipedia site for users in China circumvented the block. Within a few hours, Wikipedia was once again blocked. The prompt block on October 31 made it abundantly clear to the Chinese Wikipedia community that the block was going to be longer than expected. Since then, there have been no more efforts to appeal the blocks.

Nearly a year later, the block was partially lifted. Beginning on October 10, 2006, some parts of mainland China could access Wikipedia. On November 10, 2006, Chinese Wikipedia appeared to have been fully unblocked but was soon reblocked on November 17.

Access to Wikipedia sites was restored on June 15, 2007, with the exception of politically sensitive articles and Chinese Wikipedia. On July 25, 2007, Chinese Wikipedia was also unblocked. Within a few hours, however, all Wikipedia sites were blocked again.

On April 2, 2008, all Wikipedia sites except Chinese Wikipedia and some politically sensitive pages were unblocked. The move came two days after the International Olympic Committee (IOC) notified China that it wanted the Internet freely accessible during the Olympic Summer Games. Starting on July 3, 2008, China stopped restricting access to Chinese Wikipedia in some parts of the country. The unblock was extended to the entire country about one week before the opening of the Olympic Games.

Chinese Wikipedia offers an ideal empirical setting to study the relationship between group size and incentives to contribute. First, Wikipedia keeps the editing history of all articles, and each edit can be traced to an ID that uniquely identifies the contributor. As a result, we can accurately measure the contributions from each contributor over time.

Second, the blocks of Chinese Wikipedia present an easy way to establish a causal relationship. Changes in individual contributions might correlate with group size for

various reasons. For example, while the number of contributors at Wikipedia generally increases over time, individual contributors often contribute less over time. While this observation is consistent with the free-rider hypothesis, it might be a result of contributors having less to contribute over time, or perhaps it is simply a case of their slacking off after a "honeymoon" period.

The blocks of Chinese Wikipedia provide natural experiments to examine the impact of group size. The blocks are exogenous, as for each block there was neither a warning beforehand nor any explanation afterward. As contributors outside mainland China were not blocked, we can examine changes in their contribution levels and study how their incentives to contribute change as a result of the change in group size.

Our empirical analysis focuses on the third block, which took place in October 2005 and lasted for nearly one year. We choose to focus on the third block because it is the longest of the five blocks. In the four other instances, blocking and unblocking either happened within several days or took place together with other confounding events (e.g., the Olympic Summer Games in 2008). As a result, the impact of other blocks on individual contributions is difficult to measure. In addition, after the third block, many people believed that the ban was going to be permanent and that the subsequent unblocks were caused by glitches in the Great Firewall System. Given these expectations, contributors may not have adjusted their contribution levels accordingly.

The third block was also well publicized. The block was quickly reported in many major news agencies worldwide. In addition, in response to this block, on October 20, 2005, the home page of Chinese Wikipedia displayed a link directing any user from mainland China to a status page. Nonblocked contributors could easily learn about the block. The great publicity the block received frees us from a concern in several previous studies that individuals may be unaware of the changes in the environment and hence do not change their contribution levels. We subsequently use "before the block" and "after the block" to refer to the time periods before and after the beginning of the third block, respectively.

One potential concern is that as Wikipedia articles are durable, contributors may be unconcerned by blocks if they expect mainland Chinese to have the access again in the future. There are several reasons for even forward-looking contributors to believe that the group size of Chinese Wikipedia would be significantly reduced, particularly after the third block. The immediate reblocking during the third block and the failure to receive a response to the appeal suggested that the Chinese government had determined that Wikipedia content is inappropriate for mainland Chinese and should be blocked. Historically, after becoming targets of the Great Firewall System, many sites, such as those of BBC News, have been blocked for years. Even if the block were to be permanently lifted in the future, with the great uncertainty in accessibility, the site would no longer have the same appeal to mainland users. In addition, the Chinese government is likely to support the development of competing products by domestic firms. Google, for example, has been blocked and unblocked several times in China since 2002. As a result, Google lost significant market share to Baidu, a domestic search engine in China that is known to have a good relationship with the government.[5] Chinese Wikipedia had a similar experience. In April

[5] http://www.searchenginejournal.com/google-losing-market-share-in-china/3816/ (accessed July 2009).

2006, Baidu launched Baidu Baike, a wiki-based encyclopedia with the same functionalities as Chinese Wikipedia but with censored content. The site has not been subject to blocking. Owing to the blocks of Chinese Wikipedia, by November 2007, Baidu Baike contained more articles than any other online encyclopedias except English Wikipedia, and thus essentially replaced the role of Chinese Wikipedia in mainland China.[6] Therefore, even if contributors on Chinese Wikipedia are forward-looking and take the future group size of Chinese Wikipedia into consideration, many of them would expect Chinese Wikipedia to reach a smaller audience than it could have without the block. It is still possible, however, that some contributors expect the block to have no effect on the diffusion of Chinese Wikipedia in mainland China. In this case, the effect on these contributors should be weaker and our results, hence, would be more pronounced had all contributors believed that the reduction in group size is permanent.

II. Empirical Analysis

A. *Data*

We obtain our dataset from the Chinese Wikipedia website (http://zh.wikipedia.org/). The dataset contains the full text of all the articles and their complete editing histories. There are 196,130 articles posted between October 2002 and February 2007. To study the impact of the block, we focus on contributions of nonblocked contributors four weeks before October 19, 2005, and four weeks after October 31, 2005. We choose this relatively short time window because the group size increases again after the block. Hence, results from a big time window may not accurately reflect the impact of a reduction in group size. During this eight-week period, 9,048 new articles are initiated and 53,519 revisions are made. A total of 10,436,966 characters are added and 4,321,112 characters are deleted from these articles.[7]

Contributors are identified by their IDs if they have registered. Otherwise, they are identified by their network IP addresses at the time of connection. As the same IP address can map to multiple contributors and a contributor may not always use the same IP address, we focus on registered contributors in our analysis. We also exclude administrators and robots from the analysis, because they may exhibit different editing patterns. Among the 21,496 registered contributors at Chinese Wikipedia, there are 22 robots and 87 administrators. For each article, we collect data on the date and time of each revision, contributor ID, and the number of characters added and deleted for each revision. We then use this information to generate the contribution history of each contributor in each week.

In addition to article pages on Wikipedia, each contributor can create her own user page and/or user-talk page to facilitate communication with others.[8] Many contributors use their user pages to introduce themselves by adding such information as geographic location, contact information, photographs, and their areas of expertise and

[6] http://www.businessweek.com/globalbiz/content/nov2007/gb20071113_725400.htm (accessed July 2009).

[7] In Chinese, characters form the basic unit of meaning. Most Chinese words are formed by two or three characters.

[8] http://en.wikipedia.org/wiki/User_page (accessed July 2009).

interest. Contributors can also use user-talk pages to leave comments for each other or to discuss issues related to Wikipedia articles. Generally these user pages and user-talk pages, like Wikipedia articles, can be modified by anyone. We also obtain the editing history of each contributor in user pages and user-talk pages.

B. *Identifying Nonblocked Contributors*

To assess the block's impact on contributors outside mainland China, we first need to identify these nonblocked contributors. Wikipedia reveals neither individual contributors' geographic information nor registered contributors' network IP addresses. We rely on two information sources to identify nonblocked contributors.

First, we use each contributor's contribution history to infer whether the block affected her. We consider a contributor as nonblocked if she joins Chinese Wikipedia before the beginning of the third block and contributes at least once during any of the blocked periods. Of the 6,062 contributors who joined Wikipedia before the block, 1,623 are classified as nonblocked contributors.

Second, we rely on the encoding of the characters entered by the contributors. Chinese characters are encoded according to two major standards to be correctly displayed. For historical reasons, contributors from Taiwan, Hong Kong, and Macau adopt traditional Chinese characters, while those from mainland China, Singapore, and Malaysia use simplified Chinese characters.[9] Through analyzing the encoding, we can identify whether a contributor uses simplified or traditional Chinese. Those who use traditional Chinese to edit articles are typically outside mainland China. We analyze characters added by each contributor. Some contributors add characters in both encodings, most likely because they copy and paste content from other sites whose characters are encoded differently from the ones they are using. Hence, we consider a contributor as nonblocked if more than 50 percent of her additions are in traditional Chinese. Of the 1,207 contributors who are identified as users of traditional Chinese, 118 joined Chinese Wikipedia before the block. These numbers are smaller than the ones we obtain from the first approach, because many nonblocked contributors use simplified Chinese.[10]

We then combine the two lists of nonblocked contributors. Thirty-four contributors appear in both lists. Therefore, in total, we have 1,707 nonblocked contributors who joined Chinese Wikipedia before the block. It is possible that our classification scheme excludes those nonblocked contributors who use simplified Chinese and decided to make no contributions during the blocked periods. Our results would be strengthened had these contributors been included.

C. *Summary Statistics*

Table 1 reports summary statistics for contributions by each contributor. Individual contributions are measured by the number of characters added, the number of characters deleted, and the total number of characters added and deleted. We consider

[9] See http://en.wikipedia.org/wiki/Big5 and http://en.wikipedia.org/wiki/GB2312 (accessed May 2008).

[10] Indeed, Google reports that most searches for "Chinese Wikipedia" come from users in Singapore and Malaysia. http://www.google.com/trends?q=chinese+wikipedia (accessed November 2008).

TABLE 1—SUMMARY STATISTICS FOR CONTRIBUTIONS BY INDIVIDUAL CONTRIBUTORS

	Pre-block		Post-block		Paired *t*-test
	Mean	Standard error	Mean	Standard error	*t*-stats
Panel A. Contributions from all contributors					
Addition	1,105.53	225.67	538.01	117.00	2.89***
Deletion	528.96	167.25	174.57	38.21	2.32**
Total	1,634.49	389.77	712.58	153.13	2.69***
Panel B. Contributions from nonblocked contributors					
Addition	3,599.79	797.50	1,910.62	413.73	2.43**
Deletion	1,806.45	592.76	619.95	135.13	2.19**
Total	5,406.24	1,379.49	2,530.57	541.45	2.37**

Notes: We examine the contributions made by unique contributors before and after the block. Contributions are measured by the number of characters they added and deleted. We report the contributions made during the four weeks before October 19, 2005, in the Pre-Block column, and the contributions made during the four weeks after October 31, 2005, in the Post-Block column. In the last column, we report results from paired *t*-tests.

***Significant at the 1 percent level.

**Significant at the 5 percent level.

addition and deletion as different forms of editing because the amount of effort involved may be different. We compute each contributor's contribution level. Panel A reports the mean and standard error of contributions from both blocked and nonblocked contributors. Panel B reports results for nonblocked contributors only. We also conduct paired *t*-tests to compare contribution levels before and after the block. We find that the mean contribution level of nonblocked contributors in the pre-block period is much higher than the one for all contributors, suggesting that nonblocked contributors tend to contribute more than blocked contributors in general. We also find significant declines for all measures of individual contributions after the block.

D. *Regression Analysis*

We now examine the change in contributions of each individual contributor in a regression framework:

$$(1) \qquad Contributions_{it} = \beta_0 + \beta_1 AfterBlock_t + ControlVars_{it} + \epsilon_{it},$$

where i indexes the contributors and t indexes the weeks. The dependent variable, $Contributions_{it}$, is the weekly contributions of each nonblocked contributor to Wikipedia articles. We use the logarithms of the weekly total characters added, total characters deleted, and their sums as measures for individual contributions.[11] $AfterBlock_t$ is a dummy that equals one if the time period is after the block, and zero otherwise. We also include age, measured as the number of weeks since the contributor joined Wikipedia, as a control variable, and the square of age to control for possible nonlinear effects.

[11] We add one to these measures before taking logarithms as some numbers can be zero.

TABLE 2—DETECTING THE CHANGE IN INDIVIDUAL CONTRIBUTION LEVELS AFTER THE BLOCK

Dependent variable	Total	Addition	Delection	Total	Addition	Deletion
Model	(1)	(2)	(3)	(4)	(5)	(6)
AfterBlock	−0.463***	−0.430***	−0.293***	−0.555***	−0.498***	−0.403***
	[0.044]	[0.042]	[0.032]	[0.082]	[0.079]	[0.060]
Age	−0.025***	−0.024***	−0.013***	−0.062***	−0.063***	−0.026**
	[0.002]	[0.002]	[0.002]	[0.016]	[0.015]	[0.011]
Age^2	0.000***	0.000***	0.000***	0.001***	0.001***	0.001***
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
Observations	13,376	13,376	13,376	13,376	13,376	13,376
R^2	0.03	0.03	0.02	0.04	0.04	0.03
Specification	OLS	OLS	OLS	FE	FE	FE

Notes: Addition(Deletion) refers to the logarithm of the number of characters added to (deleted from) the articles. Total is the logarithm of the sum of Addition and Deletion. *Age* is the number of weeks since the contributor joins Wikipedia. *AfterBlock* is a dummy variable that takes the value one if the time period is after the block and zero otherwise. Heteroskedasticity-adjusted standard errors in brackets.

***Significant at the 1 percent level.

**Significant at the 5 percent level.

Table 2 reports our regression results. Model 1 uses the total number of characters added and deleted as the dependent variable. Models 2 and 3 use the number of characters added and the number of characters deleted, respectively, as the dependent variables. We then repeat the analysis, controlling for individual fixed effects, in Models 4, 5, and 6. All results show the same pattern: while the individual contribution level in general decreases at a decreasing rate with age, the block significantly reduces contribution level. In addition, the number of characters added drops more than the number of characters deleted. A back-of-the-envelope calculation based on coefficients in Model 4 suggests that individual contributions drop by 42.8 percent on average as a result of the block.[12]

We now examine the impact of social effects. If social benefits provide important motivations to contributors, we expect that contributors who derive more social benefits from contributing are affected more after the block. We use several approaches to identify these nonblocked contributors who are likely to derive great social benefits from contributing.

We first consider contributors' participation in user pages and user-talk pages. As the purpose of these pages is primarily to facilitate communication, people who care to write on these pages are the ones more active in socializing with other contributors. We compute the number of characters added and deleted by each nonblocked contributor in her user page and user-talk page before the block. We extend our baseline specification to consider the following difference-in-differences specification:

$$(2)\quad Contributions_{it} = \beta_0 + \beta_1 AfterBlock_t + \beta_2 SocialParticipation_i \times AfterBlock_t + \beta_3 SocialParticipation_i + ControlVars_{it} + \epsilon_{it},$$

[12] Following Robert Halvorsen and Raymond Palmquist (1980) and Peter E. Kennedy (1981), the percentage is calculated as $\exp(\hat{\beta}_1 - \frac{1}{2}\hat{V}(\hat{\beta}_1)) - 1$, where $\hat{\beta}_1$ is the estimate of β_1 and $\hat{V}(\hat{\beta}_1)$ is the estimate of the variance of $\hat{\beta}_1$.

Table 3—Difference-in-Differences Estimations of the Impact of the Block on Contributors with Different Levels of Social Participation

Dependent variable	Total	Addition	Deletion	Total	Addition	Deletion
Model	(1)	(2)	(3)	(4)	(5)	(6)
AfterBlock	−0.105**	−0.091**	−0.013	−0.212**	−0.170**	−0.131**
	[0.041]	[0.039]	[0.028]	[0.086]	[0.082]	[0.062]
SocialParticipation × *AfterBlock*	−0.195***	−0.186***	−0.153***	−0.190***	−0.182***	−0.150***
	[0.024]	[0.023]	[0.019]	[0.024]	[0.023]	[0.020]
SocialParticipation	0.539***	0.510***	0.372***			
	[0.018]	[0.018]	[0.015]			
Age	−0.023***	−0.022***	−0.012***	−0.062***	−0.062***	−0.026**
	[0.002]	[0.002]	[0.002]	[0.015]	[0.015]	[0.011]
Age^2	0.000***	0.000***	0.000***	0.001***	0.001***	0.001***
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
Observations	13,376	13,376	13,376	13,376	13,376	13,376
R^2	0.18	0.17	0.15	0.06	0.05	0.05
Specification	OLS	OLS	OLS	FE	FE	FE

Notes: SocialParticipation is the logarithm of the weekly average of total addition and total deletion in user pages or user-talk pages by each contributor before the block. The variable *SocialParticipation* drops in the fixed-effect specifications as its value is fixed for each contributor. Heteroskedasticity-adjusted standard errors in brackets.

***Significant at the 1 percent level.

**Significant at the 5 percent level.

where $SocialParticipation_i$ is the logarithm of the weekly average of total addition and total deletion in user pages or user-talk pages by contributor *i* before the block. $SocialParticipation_i$ thus captures how socially active a contributor is.

Table 3 shows the results. The significant positive coefficients of $SocialParticipation_i$ indicate that contributors contributing more to user pages and user-talk pages also contribute more to Wikipedia articles. The negative coefficients of the interaction variable suggest that socially active contributors are affected significantly more by the block.

The measure $SocialParticipation_i$ could also capture contributors' heterogeneity in their opportunity costs of time or their tastes for contributing. As a contributor's opportunity cost of time and taste for contributing are not likely to vary significantly during the eight-week window before and after the block, we conduct fixed-effect regressions. The fixed-effect regressions enable us to focus on the change in each individual's contribution level before and after the block, rather than the absolute level. The results are qualitatively unchanged.

Our results show that these contributors who are likely to derive a high level of social utility from contributing to Wikipedia are affected more by the block and thus suggest that social effects (e.g., social interactions) could provide an important motivation for contributing.

One disadvantage of the $SocialParticipation_i$ measure is that about 46 percent of contributors do not have user pages or user-talk pages. To better capture the heterogeneity of all contributors, we construct an alternative measure. The block may exert different impacts on contributors' collaboration networks. If social benefits dominate free-riding incentives, we expect contributors with a larger percentage of blocked collaborators to decrease their contribution levels more than those

TABLE 4—DIFFERENCE-IN-DIFFERENCES ESTIMATIONS OF THE IMPACT OF THE BLOCK ON CONTRIBUTORS WITH DIFFERENT PERCENTAGES OF COLLABORATORS BLOCKED

Dependent variable	Total	Addition	Deletion	Total	Addition	Deletion
Model	(1)	(2)	(3)	(4)	(5)	(6)
AfterBlock	−0.131*	−0.125*	−0.011	−0.163	−0.134	−0.076
	[0.068]	[0.064]	[0.048]	[0.100]	[0.096]	[0.074]
PercentageBlocked × *AfterBlock*	−1.717***	−1.568***	−1.524***	−2.506***	−2.327***	−2.090***
	[0.448]	[0.425]	[0.336]	[0.398]	[0.386]	[0.335]
PercentageBlocked	7.552***	7.058***	5.253***			
	[0.359]	[0.341]	[0.281]			
Age	−0.034***	−0.032***	−0.019***	−0.079***	−0.078***	−0.040***
	[0.002]	[0.002]	[0.002]	[0.016]	[0.015]	[0.012]
Age^2	0.000***	0.000***	0.000***	0.001***	0.001***	0.001***
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
Observations	13,376	13,376	13,376	13,376	13,376	13,376
R^2	0.11	0.11	0.09	0.05	0.04	0.04
Specification	OLS	OLS	OLS	FE	FE	FE

Notes: *PercentageBlocked* is the percentage of collaborators blocked after the third block for each contributor. The variable *PercentageBlocked* drops in the fixed-effect specifications as its value is fixed for each contributor. Heteroskedasticity-adjusted standard errors in brackets.

***Significant at the 1 percent level.

**Significant at the 5 percent level.

*Significant at the 10 percent level.

whose collaboration networks are affected less. We construct a new measure, $PercentageBlocked_i$, which measures the percentage of collaborators blocked after the block for contributor i.

We consider the following specification:

$$(3)\quad Contributions_{it} = \beta_0 + \beta_1 AfterBlock_t + \beta_2 PercentageBlocked_i \times AfterBlock_t + \beta_3 PercentageBlocked_i + ControlVars_{it} + \epsilon_{it}.$$

Table 4 reports the results. As predicted, contributors with larger percentages of blocked collaborators decrease their contribution levels more. We also conduct fixed-effect regressions and obtain similar results.

Overall, our regression results provide support that, in the case of Chinese Wikipedia, social effects dominate free-riding incentives and play a significant role in motivating contributors.

III. Robustness Checks

A natural concern is whether the results above are a consequence of seasonal effects. It could be that contributors contribute less in November than in September and early October in any year. We examine contributions of nonblocked contributors

during the same eight-week period in 2003 and 2004 and find no decline in their contributions.[13]

We are also concerned that the decrease in contributions could result from less disagreement in editing after the block. Contributors in mainland China may hold different political views from those in other regions, such as Taiwan.[14] After the block, we would expect to see less disagreement about Wikipedia articles and hence less editing. We conduct two tests. First, we focus our analysis on the creation of new articles. If the decrease is caused entirely by less disagreement rather than social factors, we would expect no decline in the level of effort in creating new articles.[15] We repeat the regression analysis, this time using the logarithm of each contributor's weekly contribution of new articles as the dependent variable. We find that after the block, contributors spend less effort creating new articles, and the decrease is mostly associated with contributors who care more about social benefits.

In the second test, we examine category information for each article. When editing, a contributor can map articles to a list of categories in the database. New categories also can be easily created. For example, an article about "auction theory" could be mapped to such categories as "applied mathematics," "economics and finance," and "game theory." We compile a list of 31,871 categories from all articles, and manually go through it.[16] In the end, we identify 2,500 categories as potentially contentious categories. We then exclude contributions to articles in these categories and repeat the analysis. We obtain similar results.

In addition, we are concerned that some users contribute because they want to influence other users by promoting the public image of a firm or a product, or advocating their own views about certain Wikipedia entries. Our test exploits Wikipedia's policy that disallows self-serving edits. A contributor can be banned temporarily or permanently if he or she contributes content or makes edits in a way that violates Wikipedia policy.[17] We repeat the analysis after excluding contributors who have been banned once or more by the administrators. We obtain similar results. Thirty-nine contributors among the 1,707 nonblocked contributors have been banned.

Our fourth concern is that there might be less need for revision after the block. After the block, fewer new articles are created because there are fewer contributors. As a result, nonblocked contributors may have fewer entries to improve. We conduct our robustness check by recomputing individual contributions, this time including only contributions to articles created before the block. We obtain similar results.

Finally, technically adept contributors in mainland China might use proxy servers to circumvent the block. As proxy servers are slow, they might contribute less than before. As our classification scheme would classify them as nonblocked contributors, we may observe a decrease in their contribution levels. This concern is

[13] We report all results for robustness checks in an online Appendix.

[14] Politically sensitive articles are often edited more frequently. For example, the six most edited articles as reported by Wikipedia are Republic of China, China, People's Republic of China, Mao Zedong, Chiang Kai-shek, and Hong Kong.

[15] In fact, the level of effort could increase since contributors could have more time creating new articles after the block.

[16] The complete list of categories can be found at http://stats.wikimedia.org/EN/CategoryOverview_ZH_Complete.htm (accessed May 2008).

[17] Some major reasons to get banned are: posting advertisements and propaganda, vandalism, disruptive editing, and personal attacks or harassment.

alleviated by a Wikipedia policy that prevents contributors from editing using open and anonymous proxies.[18] In addition, Wikipedia administrators forbid open proxies in such a way that even registered contributors cannot use open proxies to edit articles. It is still possible, however, that some contributors use closed proxy servers to access Wikipedia.

Contributors who use closed proxy servers must channel their traffic through their friends' computers outside mainland China in order to edit articles at Chinese Wikipedia. Given the complexity involved in setting up and getting access to closed proxy servers, we expect those who use closed proxy servers to be dedicated and technically adept Wikipedia contributors. Hence, these contributors are likely to have contributed significantly more than other contributors before the block or are likely to contribute to articles related to information technology. We conduct our robustness check by removing contributors whose average weekly contributions are more than four standard deviations above the mean before the block, and contributors who have contributed to categories related to information technology. We obtain similar results.

IV. Concluding Remarks

Private provision of public goods is greatly valued by society. In 2007, 60.8 million volunteers, or about 26 percent of Americans age 16 or older, performed 8.1 billion hours of unpaid service to community organizations.[19] Economic theories suggest that free-riding is a concern in these contexts. In this paper, we utilize an exogenous shock on the group size of Wikipedia contributors to study how incentives to contribute change with group size at the individual level. We find a positive relationship between group size and contributors' contribution levels. In addition, we find that contributors who are likely to care more about social benefits react to the change more strongly than those who value them less. Recognition of such social effects helps explain the existence of many public goods with a large number of contributors and the empirical observation that many contributors prefer to contribute to large online communities.

Two limitations of this study are important to emphasize. Our study examines only a particular kind of public goods. On Wikipedia, contributors are at the same time users of Wikipedia, and the more contributors, the greater the quality of the good and thus the greater the number of users and contributors. While many public goods today (e.g., public radio stations, open source software, and YouTube) share similar features, there are many other types of public goods. For example, contributors to disaster relief funds are often not at the same time beneficiaries of the funds. In addition, Wikipedia is an online public good. Whether we can generalize these results to different types of public goods provision is an interesting question for future research.

[18] See the Wikipedia article on its policy on open proxies at http://en.wikipedia.org/wiki/Wikipedia:No_open_proxies (accessed May 2008) for details.

[19] The Volunteering in America Report, released on July 28, 2008 by the Corporation for National and Community Service, available at http://www.volunteeringinamerica.gov/ (accessed November 2008).

The second limitation is related to the broad definition of "social effects." Similar to many prior studies on social effects (e.g., Charles F. Manski 1993; Andreoni and Ragan Petrie 2004), we do not distinguish different motivations that give rise to social effects. This limitation largely results from the fact that different social motivations could lead to similar behavioral patterns. Future studies could seek to understand the relative importance of different motivations.

REFERENCES

Andreoni, James. 1988a. "Privately Provided Public Goods in a Large Economy: The Limits of Altruism." *Journal of Public Economics*, 35(1): 57–73.

Andreoni, James. 1988b. "Why Free Ride? Strategies and Learning in Public Goods Experiments." *Journal of Public Economics*, 37(3): 291–304.

Andreoni, James. 1989. "Giving with Impure Altruism: Applications to Charity and Ricardian Equivalence." *Journal of Political Economy*, 97(6): 1447–58.

Andreoni, James. 1990. "Impure Altruism and Donations to Public Goods: A Theory of Warm-Glow Giving?" *Economic Journal*, 100(401): 464–77.

Andreoni, James. 2006. "Philanthropy." In *Handbook of the Economics of Giving, Altruism and Reciprocity: Applications*. Vol. 2, ed. Serge-Christophe Kolm and Jean Mercier Ythier, 1201–70. Amsterdam: North Holland.

Andreoni, James. 2007. "Giving Gifts to Groups: How Altruism Depends on the Number of Recipients." *Journal of Public Economics*, 91(9): 1731–49.

Andreoni, James, and B. Douglas Bernheim. 2009. "Social Image and the Fifty-Fifty Norm: A Theoretical and Experimental Analysis of Audience Effects." *Econometrica*, 77(5): 1607–36.

Andreoni, James, and Ragan Petrie. 2004. "Public Goods Experiments without Confidentiality: A Glimpse into Fund-Raising." *Journal of Public Economics*, 88(7–8): 1605–23.

Andreoni, James, and John Karl Scholz. 1998. "An Econometric Analysis of Charitable Giving with Interdependent Preferences." *Economic Inquiry*, 36(3): 410–28.

Bergstrom, Theodore, Lawrence Blume, and Hal Varian. 1986. "On the Private Provision of Public Goods." *Journal of Public Economics*, 29(1): 25–49.

Bilodeau, Marc, and Al Slivinski. 1996. "Toilet Cleaning and Department Chairing: Volunteering a Public Service." *Journal of Public Economics*, 59(2): 299–308.

Bliss, Christopher, and Barry Nalebuff. 1984. "Dragon-Slaying and Ballroom Dancing: The Private Supply of a Public Good." *Journal of Public Economics*, 25(1–2): 1–12.

Chamberlin, John. 1974. "Provision of Collective Goods as a Function of Group Size." *American Political Science Review*, 68(2): 707–16.

Chamberlin, John. 1978. "The Logic of Collective Action: Some Experimental Results." *Behavioral Science*, 23 (5): 441–45.

Cornes, Richard, and Todd Sandler. 1984. "Easy Riders, Joint Production, and Public Goods." *Economic Journal*, 94(375): 580–98.

Deibert, Ronald J. 2002. "Dark Guests and Great Firewalls: The Internet and Chinese Security Policy." *Journal of Social Issues*, 58(1): 143–59.

Duncan, Brian. 2004. "A Theory of Impact Philanthropy." *Journal of Public Economics*, 88(9–10): 2159–80.

Ellingsen, Tore, and Magnus Johannesson. 2008. "Pride and Prejudice: The Human Side of Incentive Theory." *American Economic Review*, 98(3): 990–1008.

Fries, Timothy L., Edward Golding, and Richard E. Romano. 1991. "Private Provision of Public Goods and the Failure of the Neutrality Property in Large Finite Economies." *International Economic Review*, 32(1): 147–57.

Goeree, Jacob K., Charles A. Holt, and Susan K. Laury. 2002. "Private Costs and Public Benefits: Unraveling the Effects of Altruism and Noisy Behavior." *Journal of Public Economics*, 83(2): 255–76.

Greenstein, Shane, and Michelle Devereux. 2009. "Wikipedia in the Spotlight." Kellogg School of Management Case 5-306-507.

Halvorsen, Robert, and Raymond Palmquist. 1980. "The Interpretation of Dummy Variables in Semilogarithmic Equations." *American Economic Review*, 70(3): 474–75.

Harbaugh, William T. 1998. "The Prestige Motive for Making Charitable Transfers." *American Economic Review*, 88(2): 277–82.

Isaac, R. Mark, and James M. Walker. 1988. "Group Size Effects in Public Goods Provision: The Voluntary Contributions Mechanism." *Quarterly Journal of Economics*, 103(1): 179–99.

Isaac, R. Mark, James M. Walker, and Arlington W. Williams. 1994. "Group Size and the Voluntary Provision of Public Goods: Experimental Evidence Utilizing Large Groups." *Journal of Public Economics*, 54(1): 1–36.

Kennedy, Peter E. 1981. "Estimation with Correctly Interpreted Dummy Variables in Semilogarithmic Equations [the Interpretation of Dummy Variables in Semilogarithmic Equations]." *American Economic Review*, 71(4): 801.

Lakhani, Karim R., and Eric von Hippel. 2003. "How Open Source Software Works: 'Free' User-to-User Assistance." *Research Policy*, 32(6): 923–43.

Ledyard, John O. 1995. "Public Goods: A Survey of Experimental Research." In *The Handbook of Experimental Economics*, ed. John H. Kagel and Alvin E. Roth, 111–94. Princeton, NJ: Princeton University Press.

Lerner, Josh, and Jean Tirole. 2002. "Some Simple Economics of Open Source." *Journal of Industrial Economics*, 50(2): 197–234.

Manski, Charles F. 1993. "Identification of Endogenous Social Effects: The Reflection Problem." *Review of Economic Studies*, 60(3): 531–42.

Marwell, Gerald, and Ruth E. Ames. 1979. "Experiments on the Provision of Public Goods. I. Resources, Interest, Group Size, and the Free-Rider Problem." *American Journal of Sociology*, 84 (6): 1335–60.

Olson, Mancur. 1965. *Logic of Collective Action: Public Goods and the Theory of Groups*. Cambridge, MA: Harvard University Press.

Palfrey, Thomas R., and Howard Rosenthal. 1984. "Participation and the Provision of Discrete Public Goods: A Strategic Analysis." *Journal of Public Economics*, 24(2): 171–93.

Rege, Mari, and Kjetil Telle. 2004. "The Impact of Social Approval and Framing on Cooperation in Public Good Situations." *Journal of Public Economics*, 88(7–8): 1625–44.

Ren, Yuqing, Robert Kraut, and Sara Kiesler. 2007. "Applying Common Identity and Bond Theory to the Design of Online Communities." *Organizational Studies*, 28 (3): 377–408.

Ribar, David C., and Mark O. Wilhelm. 2002. "Altruistic and Joy-of-Giving Motivations in Charitable Behavior." *Journal of Political Economy*, 110(2): 425–57.

Soetevent, Adriaan R. 2005. "Anonymity in Giving in a Natural Context--a Field Experiment in 30 Churches." *Journal of Public Economics*, 89(11–12): 2301–23.

Steinberg, Richard S. 1987. "Voluntary Donations and Public Expenditures in a Federal System." *American Economic Review*, 77(1): 24–36.

Sweeney, John W. 1973. "An Experimental Investigation of the Free-Rider Problem." *Economic Journal*, 2(3): 277–92.

Wikimedia Foundation. 2011. "Data Dump of Chinese Wikipedia." http://dumps.wikimedia.org/zhwik:/ (accessed May 2011).

Zhang, Xiaoquan (Michael), and Feng Zhu. 2011. "Group Size and Incentives to Contribute: A Natural Experiment at Chinese Wikipedia:Dataset" *American Economic Review.* http://www.aeaweb.org/articles.php?doi=10.1257/aer.101.4.1601

Zittrain, Jonathan, and Benjamin Edelman. 2003. "Internet Filtering in China." *IEEE Internet Computing*, 7(2): 70–77.

Web Appendix

Group Size and Incentives to Contribute: A Natural Experiment at Chinese Wikipedia

Michael Zhang and Feng Zhu

In this appendix, we first provide regression results for robustness checks discussed in Section III of the paper. We then provide results from several supplemental analyses.

1 Regression Results for Robustness Checks

- *Seasonality:* We analyze the changes in contribution levels for nonblocked contributors four weeks before October 19 and four weeks after October 31 in year 2003 and 2004 in Tables A1 and A2, respectively.
- *Controversial Articles:* First, we look at contributions in new articles. We replicate the analysis in Tables 2-4 of the paper and report the results in Tables A3-A5, this time using the logarithm of weekly contributions in new articles by each contributor as the dependent variable. We then examine category information of the articles. We repeat the regression analysis and report the results in Tables A6-A8, this time only including the contributions to articles belonging to non-contentious categories.
- *Advertisement and Indoctrination Effects:* We exclude banned IDs and repeat the analysis. We report the results in Tables A9-A11.
- *Fewer New Articles for Revision:* We conduct the robustness check by re-computing individual contributions, this time only including the contributions to articles created before the block, and report the results in Tables A12-A14.
- *Proxy Server:* We remove contributors whose average weekly contributions are more than four standard deviations above the mean before the block and contributors who have contributed to categories related to information technology. We report the results in Tables A15-A17.

2 Supplemental Analyses

2.1 Additional Figures

We provide two additional figures to demonstrate the impact of the third block. Figure A1 shows the number of new contributors at Chinese Wikipedia over time. We find that before the third block, the number of new contributors grows exponentially over time, but it drops significantly as a result of the block. We also indicate the beginning and the end of each block before 2007. In general, when a block starts, the number of new contributors decreases and, after the block is lifted, the number increases dramatically.

Figure A2 shows the search volume index for the term "Chinese Wikipedia" from Google. Google's search volume index gives how many searches have been conducted for the term, relative to the total number of searches conducted on Google over time. The search volume increased significantly right after the beginning and the end of the third block, demonstrating substantial news coverage and public attention on the block.

2.2 Additional Regression Results

To further alleviate the concern caused by the unobserved opportunity cost of time, we conducted a supplemental analysis, comparing the effect of the block on nonblocked contributors who choose to reveal more personal information to those who reveal less personal information on their user pages after controlling for the amount of effort contributors spent on their user (or user talk) pages. Those who link their contributions closely to their real-world identity are likely to care more about how their contributions to Wikipedia affect their social image. Hence, we expect them to be affected more after the block. We manually went through user pages of these nonblocked contributors and collected four pieces of information: whether the user discloses her real name, discloses her location, provides her contact information (e.g., her homepage address, email address, or instant messenger ID), or provides a link to a user talk page (which reflects a desire to interact with users or other contributors). We constructed a new variable, RevealAmount, to measure the number of pieces of information a user reveals. RevealAmount is 4 if the user reveals all four pieces of information and 0 if the user doesn't have a user page or doesn't reveal any personal information. We then used the following regression specification:

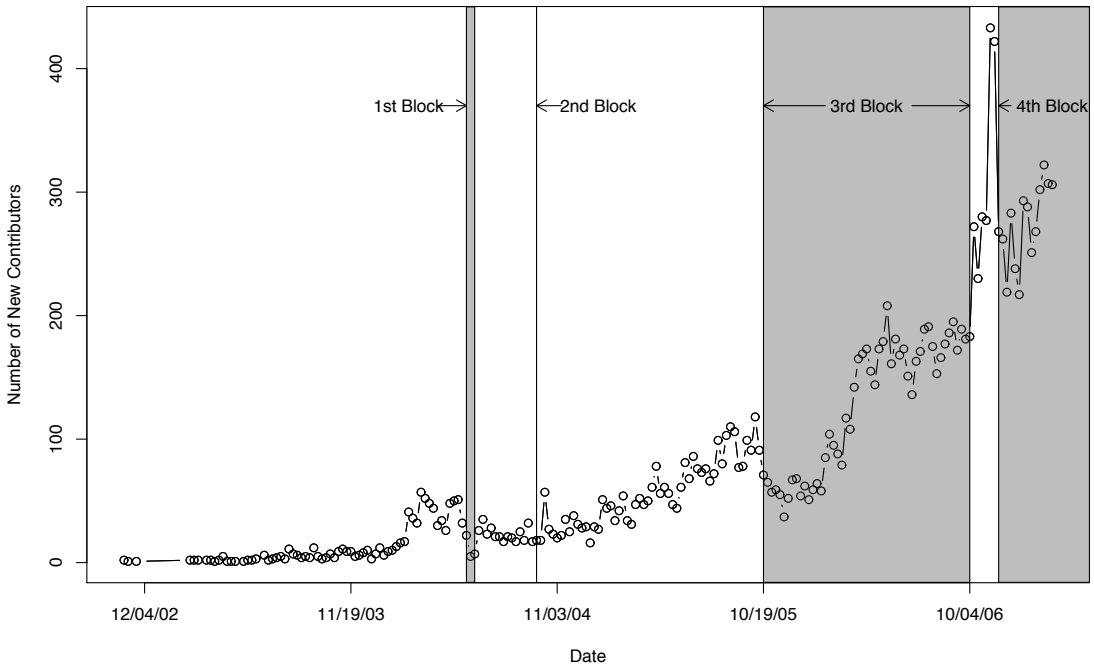
$$\begin{aligned} Contributions_{it} &= \beta_0 + \beta_1 AfterBlock_t + \beta_2 RevealAmount_i \times AfterBlock_t \\ &\quad + \beta_3 RevealAmount_i + \beta_4 SocialParticipation_i + ControlVars_{it} + \epsilon_{it}. \end{aligned} \tag{1}$$

Here, we used $SocialParticipation_i$ to control for the amount of effort a contributor spent on user pages and user talk pages, and it can also serve as a proxy to control for the opportunity cost of contributing for each individual. Both OLS and fixed-effects regressions reported in Table A18 suggest that those contributors who reveal more personal information in general contribute more and also are affected more by the block.

Our data also allow us to look at the impact of social ties among contributors. Nonblocked contributors who use Simplified Chinese are more likely to have stronger social ties with mainland Chinese and thus are more likely to care about the visibility of their contributions in mainland China and their interactions with contributors from mainland China. Thus, we expect the blocks in mainland China to reduce their incentives considerably. In contrast, contributors using Traditional Chinese are less likely to have strong ties with contributors and readers from mainland China. We thus expect their contributions to decrease less after the block. We construct a new dummy variable, $TraditionalChineseUser_i$, which is 1 if contributor i is identified as a user of Traditional Chinese, and 0 otherwise. Table A19 reports the regression results based on the following specification:

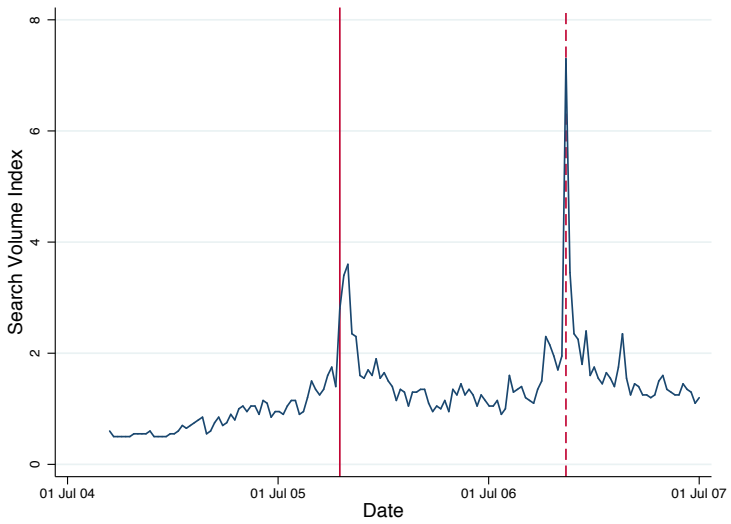
$$\begin{aligned} Contributions_{it} &= \beta_0 + \beta_1 AfterBlock_t + \beta_2 TraditionalChineseUser_i \times AfterBlock_t \\ &\quad + \beta_3 TraditionalChineseUser_i + ControlVars_{it} + \epsilon_{it}. \end{aligned} \tag{2}$$

We find that while users of Traditional Chinese contribute less on average, their contribution levels decrease much less than that of Simplified Chinese users. In other words, contributors with stronger social ties with mainland Chinese reduced their contribution levels more after the block.



Note: The shaded areas indicate the duration of the first 4 blocks.

FIGURE A1. NUMBER OF NEW CONTRIBUTORS OVER TIME



Note: The solid vertical line indicates the start of the third block and the dashed vertical line indicates the end of the third block. The y-axis (Search Volume Index) shows users' propensity to search for a certain topic on Google on a relative basis. Google does not report the number of times a term is searched, and the numbers are normalized to take the average search volume as one.

FIGURE A2. SEARCH VOLUME INDEX FOR "CHINESE WIKIPEDIA" ON GOOGLE

TABLE A1—REPLICATING TABLE 2 OF THE PAPER FOR THE SAME TIME WINDOW IN 2003

Model	(1)	(2)	(3)	(4)	(5)	(6)
Dependent Variable	Total	Addition	Deletion	Total	Addition	Deletion
AfterBlock	−0.596	−0.544	−0.277	0.083	0.195	−0.278
	[0.364]	[0.351]	[0.270]	[0.776]	[0.760]	[0.496]
Age	−0.123**	−0.122**	−0.046	−0.388***	−0.394***	−0.120
	[0.049]	[0.048]	[0.036]	[0.120]	[0.117]	[0.098]
Age^2	0.004***	0.004***	0.002***	0.008***	0.007***	0.004***
	[0.001]	[0.001]	[0.001]	[0.002]	[0.002]	[0.001]
Observations	283	283	283	283	283	283
R-squared	0.18	0.18	0.21	0.15	0.15	0.08
Number of IDs				43	43	43
Specification	OLS	OLS	OLS	FE	FE	FE

Note: For all regression tables, heteroskedasticity-adjusted standard errors in brackets; * significant at 10%; ** significant at 5%; *** significant at 1%.

TABLE A2—REPLICATING TABLE 2 OF THE PAPER FOR THE SAME TIME WINDOW IN 2004

Model	(1)	(2)	(3)	(4)	(5)	(6)
Dependent Variable	Total	Addition	Deletion	Total	Addition	Deletion
AfterBlock	0.042	0.028	0.054	0.011	−0.010	0.055
	[0.072]	[0.069]	[0.051]	[0.147]	[0.140]	[0.102]
Age	−0.084***	−0.079***	−0.052***	−0.093***	−0.082***	−0.057***
	[0.008]	[0.007]	[0.005]	[0.031]	[0.030]	[0.022]
Age^2	0.001***	0.001***	0.001***	0.001***	0.001**	0.001***
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
Observations	3,316	3,316	3,316	3,316	3,316	3,316
R-squared	0.07	0.07	0.06	0.01	0.01	0.01
Number of IDs				456	456	456
Specification	OLS	OLS	OLS	FE	FE	FE

TABLE A3—REPLICATING TABLE 2 OF THE PAPER FOR ADDITION TO NEW ARTICLES

Model	(1)	(2)
Dependent Variable	Addition to New Articles	Addition to New Articles
AfterBlock	-0.123^{***}	-0.154^{***}
	[0.027]	[0.059]
Age	-0.010^{***}	-0.025^{**}
	[0.001]	[0.011]
Age^2	0.000^{***}	0.000^{***}
	[0.000]	[0.000]
Observations	13,376	13,376
R-squared	0.01	0.01
Number of IDs		1,707
Specification	OLS	FE

TABLE A4—REPLICATING TABLE 3 OF THE PAPER FOR ADDITION TO NEW ARTICLES

Model	(1)	(2)
Dependent Variable	Addition to New Articles	Addition to New Articles
AfterBlock	0.016	−0.014
	[0.024]	[0.058]
SocialParticipation	-0.076^{***}	-0.078^{***}
× AfterBlock	[0.018]	[0.019]
SocialParticipation	0.226^{***}	
	[0.014]	
Age	-0.009^{***}	-0.025^{**}
	[0.001]	[0.011]
Age^2	0.000^{***}	0.000^{***}
	[0.000]	[0.000]
Observations	13,376	13,376
R-squared	0.08	0.01
Number of IDs		1,707
Specification	OLS	FE

TABLE A5—REPLICATING TABLE 4 OF THE PAPER FOR ADDITION TO NEW ARTICLES

Model	(1)	(2)
Dependent Variable	Addition to New Articles	Addition to New Articles
AfterBlock	-0.065^{*}	-0.069
	[0.039]	[0.066]
PercentageBlocked	-0.231	-0.541^{**}
× AfterBlock	[0.266]	[0.274]
PercentageBlocked	2.532^{***}	
	[0.214]	
Age	-0.013^{***}	-0.029^{**}
	[0.001]	[0.011]
Age^2	0.000^{***}	0.000^{***}
	[0.000]	[0.000]
Observations	13,376	13,376
R-squared	0.04	0.01
Number of IDs		1,707
Specification	OLS	FE

TABLE A6—REPLICATING TABLE 2 OF THE PAPER FOR CONTRIBUTIONS TO NON-CONTENTIOUS ARTICLES

Model	(1)	(2)	(3)	(4)	(5)	(6)
Dependent Variable	Total	Addition	Deletion	Total	Addition	Deletion
AfterBlock	-0.436^{***}	-0.403^{***}	-0.267^{***}	-0.522^{***}	-0.470^{***}	-0.383^{***}
	[0.043]	[0.041]	[0.030]	[0.081]	[0.078]	[0.060]
Age	-0.024^{***}	-0.023^{***}	-0.013^{***}	-0.063^{***}	-0.061^{***}	-0.025^{**}
	[0.002]	[0.002]	[0.002]	[0.015]	[0.015]	[0.011]
Age^2	0.000^{***}	0.000^{***}	0.000^{***}	0.001^{***}	0.001^{***}	0.001^{***}
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
Observations	13,007	13,007	13,007	13,007	13,007	13,007
R-squared	0.03	0.03	0.02	0.04	0.04	0.03
Number of IDs				1,664	1,664	1,664
Specification	OLS	OLS	OLS	FE	FE	FE

TABLE A7—REPLICATING TABLE 3 OF THE PAPER FOR CONTRIBUTIONS TO NON-CONTENTIOUS ARTICLES

Model	(1)	(2)	(3)	(4)	(5)	(6)
Dependent Variable	Total	Addition	Deletion	Total	Addition	Deletion
AfterBlock	−0.068*	−0.064*	0.015	−0.171**	−0.143*	−0.111*
	[0.041]	[0.039]	[0.028]	[0.083]	[0.080]	[0.061]
SocialParticipation	−0.198***	−0.183***	−0.152***	−0.192***	−0.178***	−0.148***
× AfterBlock	[0.024]	[0.023]	[0.019]	[0.024]	[0.023]	[0.020]
SocialParticipation	0.507***	0.477***	0.344***			
	[0.018]	[0.017]	[0.015]			
Age	−0.023***	−0.021***	−0.012***	−0.062***	−0.061***	−0.025**
	[0.002]	[0.002]	[0.002]	[0.015]	[0.015]	[0.011]
Age^2	0.000***	0.000***	0.000***	0.001***	0.001***	0.001***
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
Observations	13,007	13,007	13,007	13,007	13,007	13,007
R-squared	0.17	0.16	0.14	0.05	0.05	0.04
Number of IDs				1,664	1,664	1,664
Specification	OLS	OLS	OLS	FE	FE	FE

TABLE A8—REPLICATING TABLE 4 OF THE PAPER FOR CONTRIBUTIONS TO NON-CONTENTIOUS ARTICLES

Model	(1)	(2)	(3)	(4)	(5)	(6)
Dependent Variable	Total	Addition	Deletion	Total	Addition	Deletion
AfterBlock	−0.136**	−0.134**	−0.006	−0.153	−0.134	−0.067
	[0.063]	[0.060]	[0.044]	[0.100]	[0.095]	[0.074]
PercentageBlocked	−1.528***	−1.359***	−1.406***	−2.357***	−2.143***	−2.010***
× AfterBlock	[0.408]	[0.389]	[0.305]	[0.408]	[0.388]	[0.336]
PercentageBlocked	7.059***	6.589***	4.862***			
	[0.331]	[0.315]	[0.258]			
Age	−0.034***	−0.031***	−0.019***	−0.079***	−0.076***	−0.039***
	[0.002]	[0.002]	[0.002]	[0.016]	[0.015]	[0.012]
Age^2	0.000***	0.000***	0.000***	0.001***	0.001***	0.001***
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
Observations	13,007	13,007	13,007	13,007	13,007	13,007
R-squared	0.10	0.10	0.08	0.04	0.04	0.03
Number of IDs				1,664	1,664	1,664
Specification	OLS	OLS	OLS	FE	FE	FE

TABLE A9—REPLICATING TABLE 2 OF THE PAPER AFTER EXCLUDING BANNED IDS

Model	(1)	(2)	(3)	(4)	(5)	(6)
Dependent Variable	Total	Addition	Deletion	Total	Addition	Deletion
AfterBlock	−0.470***	−0.436***	−0.295***	−0.541***	−0.487***	−0.385***
	[0.043]	[0.041]	[0.031]	[0.082]	[0.079]	[0.060]
Age	−0.026***	−0.024***	−0.013***	−0.069***	−0.068***	−0.031***
	[0.002]	[0.002]	[0.002]	[0.015]	[0.015]	[0.011]
Age^2	0.000***	0.000***	0.000***	0.001***	0.001***	0.001***
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
Observations	13,066	13,066	13,066	13,066	13,066	13,066
R-squared	0.04	0.04	0.02	0.04	0.04	0.03
Number of IDs				1,668	1,668	1,668
Specification	OLS	OLS	OLS	FE	FE	FE

TABLE A10—REPLICATING TABLE 3 OF THE PAPER AFTER EXCLUDING BANNED IDS

Model	(1)	(2)	(3)	(4)	(5)	(6)
Dependent Variable	Total	Addition	Deletion	Total	Addition	Deletion
AfterBlock	−0.137***	−0.119***	−0.038	−0.219**	−0.178**	−0.134**
	[0.041]	[0.038]	[0.027]	[0.086]	[0.083]	[0.062]
SocialParticipation	−0.185***	−0.176***	−0.142***	−0.179***	−0.172***	−0.140***
× AfterBlock	[0.025]	[0.024]	[0.019]	[0.024]	[0.023]	[0.019]
SocialParticipation	0.530***	0.501***	0.361***			
	[0.018]	[0.018]	[0.015]			
Age	−0.024***	−0.022***	−0.012***	−0.068***	−0.067***	−0.031***
	[0.002]	[0.002]	[0.002]	[0.015]	[0.015]	[0.011]
Age^2	0.000***	0.000***	0.000***	0.001***	0.001***	0.001***
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
Observations	13,066	13,066	13,066	13,066	13,066	13,066
R-squared	0.18	0.18	0.15	0.06	0.05	0.05
Number of IDs				1,668	1,668	1,668
Specification	OLS	OLS	OLS	FE	FE	FE

TABLE A11—REPLICATING TABLE 4 OF THE PAPER AFTER EXCLUDING BANNED IDS

Model	(1)	(2)	(3)	(4)	(5)	(6)
Dependent Variable	Total	Addition	Deletion	Total	Addition	Deletion
AfterBlock	−0.147**	−0.140**	−0.020	−0.152	−0.127	−0.061
	[0.067]	[0.064]	[0.047]	[0.100]	[0.096]	[0.074]
PercentBlocked	−1.666***	−1.518***	−1.488***	−2.486***	−2.303***	−2.072***
× AfterBlock	[0.446]	[0.424]	[0.334]	[0.399]	[0.386]	[0.334]
PercentBlocked	7.451***	6.962***	5.182***			
	[0.360]	[0.341]	[0.281]			
Age	−0.035***	−0.033***	−0.020***	−0.085***	−0.083***	−0.045***
	[0.002]	[0.002]	[0.002]	[0.016]	[0.015]	[0.012]
Age^2	0.000***	0.000***	0.000***	0.001***	0.001***	0.001***
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
Observations	13,066	13,066	13,066	13,066	13,066	13,066
R-squared	0.12	0.11	0.09	0.05	0.05	0.04
Number of IDs				1,668	1,668	1,668
Specification	OLS	OLS	OLS	FE	FE	FE

TABLE A12—REPLICATING TABLE 2 OF THE PAPER FOR CONTRIBUTIONS TO ARTICLES CREATED BEFORE THE BLOCK

Model	(1)	(2)	(3)	(4)	(5)	(6)
Dependent Variable	Total	Addition	Deletion	Total	Addition	Deletion
AfterBlock	−0.580***	−0.554***	−0.355***	−0.659***	−0.609***	−0.453***
	[0.042]	[0.040]	[0.030]	[0.080]	[0.077]	[0.059]
Age	−0.025***	−0.024***	−0.013***	−0.071***	−0.072***	−0.031***
	[0.002]	[0.002]	[0.002]	[0.015]	[0.014]	[0.011]
Age^2	0.000***	0.000***	0.000***	0.001***	0.001***	0.001***
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
Observations	13,376	13,376	13,376	13,376	13,376	13,376
R-squared	0.04	0.04	0.03	0.06	0.06	0.04
Number of IDs				1,707	1,707	1,707
Specification	OLS	OLS	OLS	FE	FE	FE

TABLE A13—REPLICATING TABLE 3 OF THE PAPER FOR CONTRIBUTIONS TO ARTICLES CREATED BEFORE THE BLOCK

Model	(1)	(2)	(3)	(4)	(5)	(6)
Dependent Variable	Total	Addition	Deletion	Total	Addition	Deletion
AfterBlock	−0.141***	−0.126***	−0.022	−0.234***	−0.191**	−0.127**
	[0.040]	[0.037]	[0.027]	[0.083]	[0.079]	[0.060]
SocialParticipation	−0.241***	−0.235***	−0.183***	−0.236***	−0.232***	−0.180***
× AfterBlock	[0.023]	[0.022]	[0.019]	[0.023]	[0.022]	[0.019]
SocialParticipation	0.539***	0.510***	0.372***			
	[0.018]	[0.018]	[0.015]			
Age	−0.023***	−0.022***	−0.012***	−0.071***	−0.071***	−0.031***
	[0.002]	[0.002]	[0.002]	[0.015]	[0.014]	[0.011]
Age^2	0.000***	0.000***	0.000***	0.001***	0.001***	0.001***
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
Observations	13,376	13,376	13,376	13,376	13,376	13,376
R-squared	0.18	0.18	0.15	0.08	0.08	0.06
Number of IDs				1,707	1,707	1,707
Specification	OLS	OLS	OLS	FE	FE	FE

TABLE A14—REPLICATING TABLE 4 OF THE PAPER FOR CONTRIBUTIONS TO ARTICLES CREATED BEFORE THE BLOCK

Model	(1)	(2)	(3)	(4)	(5)	(6)
Dependent Variable	Total	Addition	Deletion	Total	Addition	Deletion
AfterBlock	−0.198***	−0.191***	−0.034	−0.192**	−0.161*	−0.075
	[0.065]	[0.061]	[0.046]	[0.097]	[0.092]	[0.073]
PercentageBlocked	−2.050***	−1.948***	−1.781***	−2.989***	−2.867***	−2.412***
× AfterBlock	[0.431]	[0.406]	[0.326]	[0.404]	[0.390]	[0.347]
PercentageBlocked	7.463***	6.963***	5.209***			
	[0.358]	[0.340]	[0.280]			
Age	−0.034***	−0.032***	−0.019***	−0.091***	−0.091***	−0.047***
	[0.002]	[0.002]	[0.002]	[0.015]	[0.015]	[0.012]
Age^2	0.000***	0.000***	0.000***	0.001***	0.001***	0.001***
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
Observations	13,376	13,376	13,376	13,376	13,376	13,376
R-squared	0.12	0.11	0.09	0.07	0.07	0.05
Number of IDs				1,707	1,707	1,707
Specification	OLS	OLS	OLS	FE	FE	FE

TABLE A15—REPLICATING TABLE 2 OF THE PAPER AFTER REMOVING DEDICATED OR TECHNICALLY ADEPT CONTRIBUTORS

Model	(1)	(2)	(3)	(4)	(5)	(6)
Dependent Variable	Total	Addition	Deletion	Total	Addition	Deletion
AfterBlock	-0.375^{***}	-0.350^{***}	-0.218^{***}	-0.505^{***}	-0.459^{***}	-0.314^{***}
	[0.039]	[0.037]	[0.026]	[0.097]	[0.093]	[0.063]
Age	-0.031^{***}	-0.030^{***}	-0.016^{***}	-0.069^{***}	-0.067^{***}	-0.031^{***}
	[0.002]	[0.002]	[0.002]	[0.017]	[0.017]	[0.011]
Age^2	0.000^{***}	0.000^{***}	0.000^{***}	0.001^{***}	0.001^{***}	0.001^{***}
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
Observations	8,411	8,411	8,411	8,411	8,411	8,411
R-squared	0.06	0.05	0.04	0.05	0.04	0.03
Number of IDs				1,079	1,079	1,079
Specification	OLS	OLS	OLS	FE	FE	FE

TABLE A16—REPLICATING TABLE 3 OF THE PAPER AFTER REMOVING DEDICATED OR TECHNICALLY ADEPT CONTRIBUTORS

Model	(1)	(2)	(3)	(4)	(5)	(6)
Dependent Variable	Total	Addition	Deletion	Total	Addition	Deletion
AfterBlock	-0.147^{***}	-0.131^{***}	-0.060^{**}	-0.280^{***}	-0.240^{**}	-0.156^{**}
	[0.041]	[0.038]	[0.026]	[0.098]	[0.094]	[0.063]
SocialParticipation	-0.178^{***}	-0.171^{***}	-0.122^{***}	-0.170^{***}	-0.165^{***}	-0.119^{***}
× AfterBlock	[0.030]	[0.029]	[0.022]	[0.033]	[0.033]	[0.026]
SocialParticipation	0.319^{***}	0.304^{***}	0.196^{***}			
	[0.025]	[0.024]	[0.018]			
Age	-0.030^{***}	-0.028^{***}	-0.015^{***}	-0.065^{***}	-0.063^{***}	-0.029^{**}
	[0.002]	[0.002]	[0.002]	[0.017]	[0.017]	[0.011]
Age^2	0.000^{***}	0.000^{***}	0.000^{***}	0.001^{***}	0.001^{***}	0.001^{***}
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
Observations	8,411	8,411	8,411	8,411	8,411	8,411
R-squared	0.11	0.11	0.08	0.06	0.06	0.04
Number of IDs				1,079	1,079	1,079
Specification	OLS	OLS	OLS	FE	FE	FE

TABLE A17—REPLICATING TABLE 4 OF THE PAPER AFTER REMOVING DEDICATED OR TECHNICALLY ADEPT CONTRIBUTORS

Model	(1)	(2)	(3)	(4)	(5)	(6)
Dependent Variable	Total	Addition	Deletion	Total	Addition	Deletion
AfterBlock	-0.189^{***}	-0.178^{***}	-0.081^{**}	-0.229^{**}	-0.200^{*}	-0.127^{*}
	[0.058]	[0.055]	[0.037]	[0.114]	[0.109]	[0.075]
PercentBlocked	-1.285^{***}	-1.191^{***}	-0.963^{***}	-2.138^{***}	-2.004^{***}	-1.447^{***}
× AfterBlock	[0.400]	[0.379]	[0.263]	[0.464]	[0.443]	[0.338]
PercentBlocked	3.009^{***}	2.780^{***}	1.852^{***}			
	[0.340]	[0.323]	[0.231]			
Age	-0.035^{***}	-0.033^{***}	-0.018^{***}	-0.081^{***}	-0.078^{***}	-0.039^{***}
	[0.003]	[0.002]	[0.002]	[0.018]	[0.017]	[0.012]
Age^2	0.000^{***}	0.000^{***}	0.000^{***}	0.001^{***}	0.001^{***}	0.001^{***}
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
Observations	8,411	8,411	8,411	8,411	8,411	8,411
R-squared	0.07	0.07	0.05	0.05	0.05	0.04
Number of IDs				1,079	1,079	1,079
Specification	OLS	OLS	OLS	FE	FE	FE

TABLE A18—DIFFERENCES-IN-DIFFERENCES ESTIMATIONS OF THE IMPACT OF THE BLOCK ON CONTRIBUTORS REVEALING DIFFERENT AMOUNT OF PERSONAL INFORMATION

Model	(1)	(2)	(3)	(4)	(5)	(6)
Dependent Variable	Total	Addition	Deletion	Total	Addition	Deletion
AfterBlock	-0.388^{***}	-0.361^{***}	-0.235^{***}	-0.471^{***}	-0.419^{***}	-0.339^{***}
	[0.045]	[0.043]	[0.032]	[0.085]	[0.082]	[0.062]
RevealAmount	-0.129^{**}	-0.119^{**}	-0.100^{**}	-0.155^{***}	-0.144^{**}	-0.117^{**}
× AfterBlock	[0.057]	[0.055]	[0.044]	[0.059]	[0.057]	[0.048]
RevealAmount	0.140^{***}	0.131^{***}	0.117^{***}			
	[0.045]	[0.043]	[0.035]			
SocialParticipation	0.429^{***}	0.405^{***}	0.284^{***}			
	[0.013]	[0.013]	[0.010]			
Age	-0.024^{***}	-0.023^{***}	-0.012^{***}	-0.065^{***}	-0.065^{***}	-0.028^{**}
	[0.002]	[0.002]	[0.002]	[0.016]	[0.015]	[0.011]
Age^2	0.000^{***}	0.000^{***}	0.000^{***}	0.001^{***}	0.001^{***}	0.001^{***}
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
Observations	13,376	13,376	13,376	13,376	13,376	13,376
R-squared	0.17	0.17	0.14	0.04	0.04	0.03
Number of IDs				1,707	1,707	1,707
Specification	OLS	OLS	OLS	FE	FE	FE

TABLE A19—DIFFERENCES-IN-DIFFERENCES ESTIMATIONS OF THE IMPACT OF THE BLOCK ON CONTRIBUTORS USING DIFFERENT CHARACTER ENCODINGS

Model	(1)	(2)	(3)	(4)	(5)	(6)
Dependent Variable	Total	Addition	Deletion	Total	Addition	Deletion
AfterBlock	-0.492^{***}	-0.458^{***}	-0.314^{***}	-0.585^{***}	-0.526^{***}	-0.424^{***}
	[0.046]	[0.044]	[0.034]	[0.083]	[0.080]	[0.061]
TraditionalChineseUser	0.403^{***}	0.384^{***}	0.292^{***}	0.449^{***}	0.426^{***}	0.324^{***}
× AfterBlock	[0.079]	[0.075]	[0.051]	[0.076]	[0.073]	[0.052]
TraditionalChineseUser	-1.290^{***}	-1.217^{***}	-0.805^{***}			
	[0.064]	[0.061]	[0.041]			
Age	-0.024^{***}	-0.023^{***}	-0.012^{***}	-0.063^{***}	-0.064^{***}	-0.027^{**}
	[0.002]	[0.002]	[0.002]	[0.016]	[0.015]	[0.011]
Age^2	0.000^{***}	0.000^{***}	0.000^{***}	0.001^{***}	0.001^{***}	0.001^{***}
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
Observations	13,376	13,376	13,376	13,376	13,376	13,376
R-squared	0.05	0.05	0.03	0.04	0.04	0.03
Number of IDs				1,707	1,707	1,707
Specification	OLS	OLS	OLS	FE	FE	FE