Web Appendices for
"Understanding Users' Content Contribution Behavior When Knowledge Can be Priced"

September 6, 2024

Contents

A Additional Information on Zhihu **3**

A.1 The User Profile Page 3

A.2 The Interface of Live Event 4

A.3 Supplementary Information on the Program Announcement 5

B Average Treatment Effect of Zhihu Live Program **7**

B.1 Distribution of Weekly Number of Answers in User-Week Panel 7

B.2 Validating the DiD Model 8

B.2.1 Relative Time Model and Falsification Exercise 8

B.2.2 Selection on Observables 9

B.2.3 Selection on Unobservables 12

B.3 Additional Robustness Checks 15

B.3.1 Controlling for User Tenure 15

B.3.2 Removal of Inactive Nonparticipants 16

B.3.3 Exclusion of Outliers 16

B.3.4 Additional Functional Forms 16

C Mechanism **18**

C.1 Extract Question and Answer Characteristics 18

C.1.1 Classify Questions and Answers by Topical Category 18

C.1.2 Identify Open-ended Questions 20

C.2 An Example of A Post-event Answer 22

C.3 Temporal Distance Effects across Users 22

A Additional Information on Zhihu

A.1 The User Profile Page

Figure W1: An Example of A User's Profile Page

A.2 The Interface of Live Event

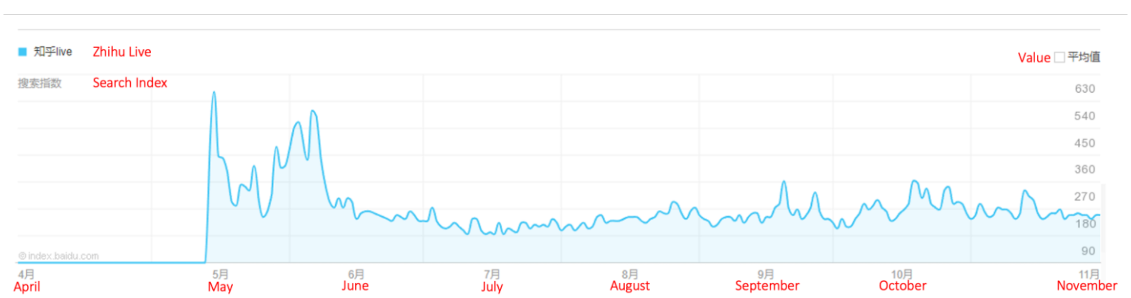
Figure W2: An Example of A Live Event

A.3 Supplementary Information on the Program Announcement

We analyze online users' search interests in "Zhihu Live" using the Baidu Index (http://index.baidu.com). The Baidu Index provides time series data similar to Google Trends for tracking user interest levels in specific keywords. Figure W3 presents the search interests in Zhihu Live from April 2016 to November 2016. The search volume for this keyword is zero before mid-May, peaks between mid-May and mid-June, and remains stable afterward. This trend indicates that the public was not aware of the Zhihu Live program before its official announcement.

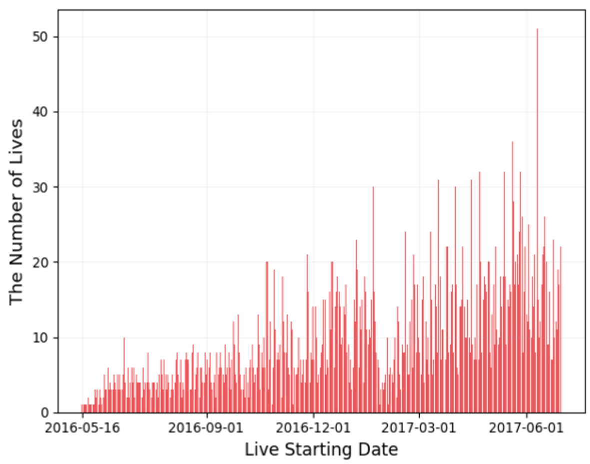
Although Zhihu maintained confidentiality about the the program's development, it is possible that the information was shared with some influential users who were invited to prepare for paid content before the official announcement. If this is the case, we would expect a surge in the number of Live events shortly after the program announcement, with some events having been created even before the announcement date. To investigate this hypothesis, we examine the distribution of the daily number of Live events in Figure W4. We find that very few events are streamed soon after the announcement and none are created before it. Moreover, the host of the first Live event, as documented in Figure W5, received her invitation from Zhihu on May 14, 2016 and hosted the event on May 16. This evidence collectively enhances our confidence that the program announcement is exogenous to users.

Figure W3: Baidu Index of "Zhihu Live"

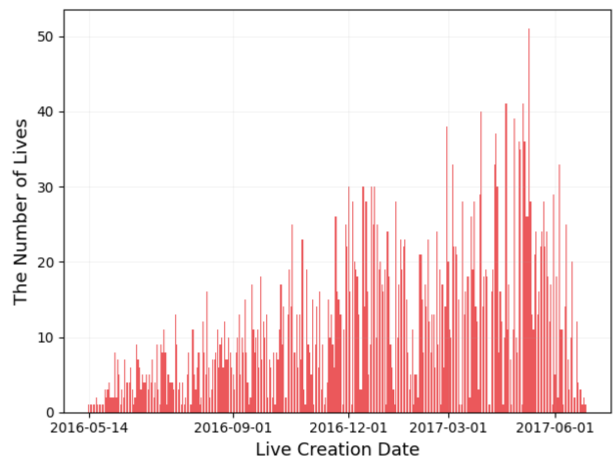


Note: This figure plots the time series of the Baidu Index of Zhihu Live from April 2016 to November 2016.

Figure W4: Daily Number of Live Events



(a) Daily Number of New Live Events



(b) Daily Number of Newly Created Live Events

Note. The x-axis is the number of weeks relative to the week of Live event start or creation.

Figure W5: The First Live Event on Zhihu

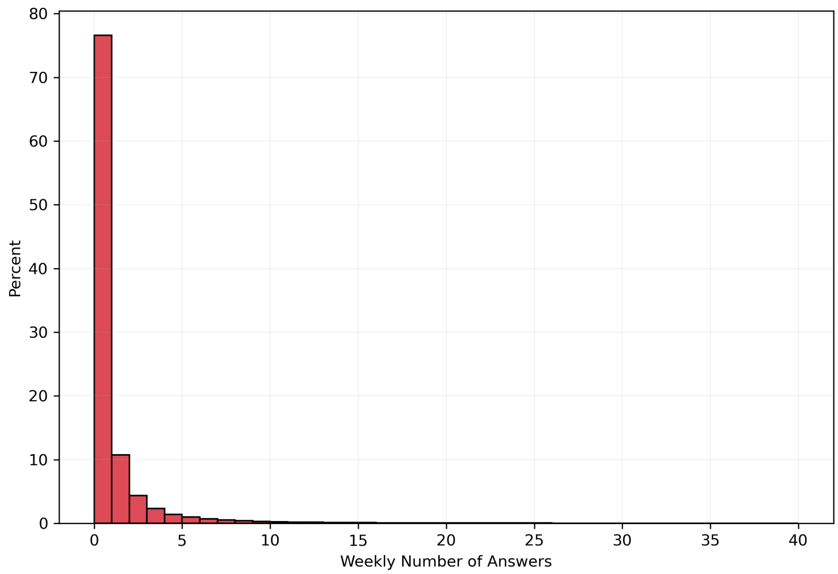
(a) The First Live Event

(b) An Article Posted by the Event Host

B Average Treatment Effect of Zhihu Live Program

B.1 Distribution of Weekly Number of Answers in User-Week Panel

Figure W6: Distribution of Weekly Number of Answers in User-Week Panel



B.2 Validating the DiD Model

B.2.1 Relative Time Model and Falsification Exercise

Table W1: Relative Time Model and Falsification Test of Pretreatment Trends

		Falsification Test				
DV: $Answer_{it}$	Relative Time Model	Placebo Treatment 2015-10-17	Placebo Treatment 2015-10-24	Placebo Treatment 2015-10-31	Placebo Treatment 2015-11-07	Placebo Treatment 2015-11-14
$EP_i \times After_t^{-16^+}$	-0.207^{***}					
	(0.075)					
$EP_i \times After_t^{-15}$	-0.030					
	(0.087)					
$EP_i \times After_t^{-14}$	-0.087					
	(0.078)					
$EP_i \times After_t^{-13}$	-0.021					
	(0.093)					
$EP_i \times After_t^{-12}$	-0.010					
	(0.105)					
$EP_i \times After_t^{-11}$	-0.172					
	(0.106)					
$EP_i \times After_t^{-10}$	-0.171^{**}					
	(0.071)					
$EP_i \times After_t^{-9}$	-0.078					
	(0.069)					
$EP_i \times After_t^{-8}$	-0.107					
	(0.072)					
$EP_i \times After_t^{-7}$	-0.042					
	(0.066)					
$EP_i \times After_t^{-6}$	-0.085					
	(0.074)					
$EP_i \times After_t^{-5}$	-0.039					
	(0.064)					
$EP_i \times After_t^{-4}$	-0.059					
	(0.061)					
$EP_i \times After_t^{-3}$	0.074					
	(0.081)					
$EP_i \times After_t^{-2}$	0.029					
	(0.073)					
$EP_i \times After_t^{-1}$	omitted					
$EP_i \times After_t$	0.453^{***}	-0.010	-0.001	0.006	0.013	0.020
	(0.056)	(0.061)	(0.062)	(0.061)	(0.062)	(0.062)
Week fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
User fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Num. users	50,961	49,150	49,150	49,150	49,150	49,150
Num. obs.	5,962,437	2,850,700	2,850,700	2,850,700	2,850,700	2,850,700
Log Pseudo-likelihood	-6.730×10^6	-3.984×10^6	-3.984×10^6	-3.984×10^6	-3.984×10^6	-3.984×10^6

Note. Robust standard errors clustered at the user level are in the parentheses. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

B.2.2 Selection on Observables

Table W2: Propensity Score Matching: t-Test for Covariates of Matched User Pairs.

	Before matching			After matching		
Covariates	Nonparticipant	Participant	t-Test, p-value	Nonparticipant	Participant	t-Test, p-value
Education	0.04	0.12	0.000	0.13	0.12	0.312
OtherAccount	0.10	0.38	0.000	0.41	0.38	0.113
Founder	0.02	0.07	0.000	0.06	0.07	0.548
Profession	0.04	0.22	0.000	0.22	0.22	0.880
Career	0.02	0.04	0.000	0.04	0.03	0.822
Monetization	0.01	0.08	0.000	0.07	0.08	0.576
Gender	0.70	0.82	0.000	0.86	0.83	0.177
Email	0.96	0.92	0.000	0.92	0.92	0.646
Phone	0.48	0.98	0.000	0.98	0.98	0.868
Tenure	141.06	138.99	0.298	135.60	138.76	0.311
Follower	1315.66	8373.76	0.000	7696.72	8212.47	0.342
Followee	149.83	134.66	0.150	141.67	134.70	0.566
Answer	138.21	186.56	0.000	204.50	185.92	0.258
Question	8.64	13.49	0.000	12.63	13.55	0.724
Upvote	3725.48	13162.64	0.000	19549.90	20825.49	0.665
Downvote	254.85	914.04	0.000	891.45	914.04	0.888
Thank	1063.58	5765.01	0.000	5250.45	5765.01	0.486
Unhelpful	267.74	1150.44	0.000	1044.55	1150.44	0.491
UpvoteClicked	722.84	709.05	0.789	744.21	711.87	0.654
DownvoteClicked	112.45	217.37	0.000	221.03	217.76	0.922
ThankClicked	173.99	217.56	0.212	218.57	218.50	1.000
UnhelpfulClicked	79.87	134.01	0.000	136.62	134.25	0.924
QuestionSubscribed	395.35	368.77	0.415	395.77	369.17	0.363
TopicSubscribed	48.52	36.59	0.003	37.67	36.59	0.679
$AvgAnswer^{pre}$	1.22	1.25	0.763	1.40	1.25	0.177
$AvgUpvote^{pre}$	51.99	109.59	0.000	120.99	108.90	0.375

Table W3: Average Treatment Effect of Zhihu Live Program on Answer Contributions on Matched Sample.

	Dependent Variable: $Answer_{it}$	
	Program announcement as the breakpoint	Participant actual enrollment as breakpoints
$EP \times After_t$	0.406^{***}	
	(0.081)	
$EP \times After_{it}$		0.270^{***}
		(0.063)
Week fixed effects	Yes	Yes
User fixed effects	Yes	Yes
Num. users	2,290	2,290
Num. obs.	267,930	267,930
Log Pseudo-likelihood	-3.660×10^5	-3.670×10^5

Note. Robust standard errors clustered at the user level are in the parentheses. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

Table W4: Relative Time Model and Falsification Test of Pretreatment Trends on Matched Sample.

		Falsification Test				
DV: $Answer_{it}$	Relative Time Model	Placebo Treatment 2015-10-17	Placebo Treatment 2015-10-24	Placebo Treatment 2015-10-31	Placebo Treatment 2015-11-07	Placebo Treatment 2015-11-14
$EP_i \times After_t^{-16^+}$	-0.029					
	(0.101)					
$EP_i \times After_t^{-15}$	0.033					
	(0.115)					
$EP_i \times After_t^{-14}$	0.031					
	(0.105)					
$EP_i \times After_t^{-13}$	0.142					
	(0.117)					
$EP_i \times After_t^{-12}$	0.078					
	(0.132)					
$EP_i \times After_t^{-11}$	-0.030					
	(0.132)					
$EP_i \times After_t^{-10}$	-0.074					
	(0.104)					
$EP_i \times After_t^{-9}$	0.003					
	(0.101)					
$EP_i \times After_t^{-8}$	-0.078					
	(0.104)					
$EP_i \times After_t^{-7}$	0.014					
	(0.099)					
$EP_i \times After_t^{-6}$	0.058					
	(0.104)					
$EP_i \times After_t^{-5}$	0.009					
	(0.100)					
$EP_i \times After_t^{-4}$	0.076					
	(0.092)					
$EP_i \times After_t^{-3}$	0.141					
	(0.099)					
$EP_i \times After_t^{-2}$	0.050					
	(0.103)					
$EP_i \times After_t^{-1}$	omitted					
$EP_i \times After_t$	0.381^{***}	-0.100	-0.086	-0.075	-0.068	-0.066
	(0.086)	(0.079)	(0.079)	(0.079)	(0.079)	(0.080)
Week fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
User fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Num. users	2,290	2,089	2,089	2,089	2,089	2,089
Num. obs.	267,930	121,162	121,162	121,162	121,162	121,162
Log Pseudo-likelihood	-3.660×10^5	-1.814×10^5	-1.815×10^5	-1.815×10^5	-1.815×10^5	-1.815×10^5

Note. Robust standard errors clustered at the user level are in the parentheses. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

B.2.3 Selection on Unobservables

The key assumption of matching is that the selection process is fully accounted for by observable variables. However, hidden bias may arise if unobserved variables simultaneously affect both the treatment assignment and the outcome variables. To evaluate the robustness of our estimation against potential hidden bias, we employ the widely adopted Rosenbaum bounds test for sensitivity analysis (Rosenbaum 2002). Rosenbaum bounds provide critical information on the level of uncertainty in matching estimators by quantifying the magnitude of influence that unobservables would need to exert to undermine the conclusions drawn from a matching analysis (Rosenbaum 2002, DiPrete and Gangl 2004).

Consider a matched pair of users with identical observed covariates X. Let P_T represent the propensity of the participant to enroll into the Zhihu Live program, specified as a function of observed covariates,

$$P_T = X_T \beta + u_T, \tag{1}$$

where T indexes the participants, u_T represents the effect of unobservables on the decision to select into the program. If unobservables have no influence on the selection process, $u_T = 0$. Similarly, let P_C denote the propensity of the nonparticipant in the same matched pair to enroll. The odds that each user receives treatment are then represented by $P_T/(1-P_T)$ and $P_C/(1-P_C)$, and the odds ratio is $\frac{P_T/(1-P_T)}{P_C/(1-P_C)} = \frac{P_T(1-P_C)}{(1-P_T)P_C}$. In the absence of unobserved selection, the odds-ratio of participants to nonparticipants is 1. The intuition is that users who are similar in terms of observed covariates should have an equal probability of enrolling in the program if unobserved factors do not play a role.

However, the matched users may nonetheless differ in their odds of selecting into the program due to the effect of unobserved selection. Rosenbaum bounds impose a constraint on the extent to which unobserved selection can alter the odds ratio by a parameter (Γ) such that

$$\frac{1}{\Gamma} \leq \frac{P_T(1-P_C)}{(1-P_T)P_C} \leq \Gamma \tag{2}$$
$$\Rightarrow \frac{1}{1+\Gamma} \leq \frac{P_T}{(P_T+P_C)} \leq \frac{\Gamma}{1+\Gamma}\; for\; all\; X_T = X_C.$$

Γ serves as an approximate measure of the bias level introduced in our treatment effect because of unobserved selection or the degree of deviation from randomized treatment assignment (Rosenbaum et al. 2005). If $\Gamma = 1$, both matched users have the same probability of enrolling into the Zhihu Live program and the study is free of hidden bias. If $\Gamma = 2$, one user might be twice as likely as another to enroll because of unobserved pre-treatment differences. The sensitivity analysis varies the value of Γ to examine how much hidden bias can exist before the qualitative conclusions of the study are affected. A study is highly sensitive to hidden bias if the conclusions change with Γ slightly greater than 1, and it is considered insensitive if the conclusions withstand much larger values of Γ (Rosenbaum et al. 2005).

Rosenbaum bounds typically employ the Wilcoxon signed-rank statistic or McNemar's statistic for matched pairs to test the null hypothesis that outcomes are identical between matched observations from treated and control individuals. In our study, we utilize the Wilcoxon signed-rank (W) statistic. If we have a sufficiently large number of observations to match participants and non-participants based on the propensity score, the test statistics asymptotically approximate a normal distribution, allowing for the use a p-value for "W" under the null.[1] The parameter Γ quantifies the degree to which the probability of treatment assignment deviates from the value under Wilcoxon null distributions, which predict a probability of 0.5, conditional on observed covariates. The ranking of each of the $N = \{1, \ldots, 1153\}$ matched pairs is now included with probability for the upper bound and lower bound distributions of $\frac{\Gamma}{1+\Gamma}$ and $1/(1+\Gamma)$, respectively. Given the W-statistics calculated from the data, we can calculate the probability that this value of W was drawn from null distributions of the upper and lower bounds respectively (columns sig- and sig+ in Table W5). If

[1] See Rosenbaum (2002) for further details on the applicability of the asymptotic approximation in applied settings.

the resulting p-values reject the null hypothesis that there is no difference in answer contributions between (matched) participants and nonparticipants, our results are robust to unobserved selection for the the specified level of Γ.

Table W5 shows the results of the Rosenbaum bounds test. As our main DiD analysis identifies a positive effect of the paid feature on answer contributions, we are primarily concerned with potential upward (positive) than downward (negative) bias in the DiD estimator. Therefore, we are mostly interested in the sig+ columns in Table W5. We find that the estimate is significant for $\Gamma <= 1.8$. This suggests that unobservables would need to increase the odds of selection into the program by at least 80% to overturn the positive estimated treatment effect on answer contributions. The results of our sensitivity analysis are similar to the results obtained in the extant literature (DiPrete and Gangl 2004, Sun and Zhu 2013, Oestreicher-Singer and Zalmanson 2013, Manchanda et al. 2015, Zhang et al. 2022), which report results with scales ranging from 1.2 to 1.8. This suggests that our study is robust to the hypothetical unobserved factors that may affect the treatment likelihood.

Table W5: Rosenbaum Bounds Test Results

Sensitivity parameter Γ	Significance level	
	sig-	sig+
1	0	0
1.1	0	0
1.2	0	0
1.3	0	0
1.4	0	0
1.5	0	0
1.6	0	0.001
1.7	0	0.014
1.8	0	0.080
1.9	0	0.255
2.0	0	0.519

B.3 Additional Robustness Checks

B.3.1 Controlling for User Tenure

Our previous findings suggest that users' incentives for contributions tend to decrease over time. Therefore, if nonparticipants are "older" at the time of the program announcement, we may overestimate the program effect by comparing the trends of "young" participants to the declining trends of nonparticipants whose incentives for contributions are starting to diminish. However, as shown in Table 4, nonparticipants join Zhihu on average just two weeks earlier than participants. Despite this minor difference, we add user tenure as an additional control variable in our main specification. Column (1) in Table W6 presents the estimation results. We find that the inclusion of user tenure does not change the point estimate of the coefficient on $EP_i \times After_t$. Therefore, we conclude that the two-week difference in user tenure between the two groups is not the driver of the observed program effect.

Table W6: Average Treatment Effect of Zhihu Live Program on Answer Contributions: Robustness to the Inclusion of Additional Control and Exclusion of Inactive Accounts and Outliers

	Dependent Variable: $Answer_{it}$				
	(1)	(2)	(3)	(4)	(5)
$EP_i \times After_t$	0.631^{***}	0.579^{***}	0.631^{***}	0.630^{***}	0.766^{***}
	(0.054)	(0.054)	(0.054)	(0.054)	(0.048)
$Tenure_{it}$	0.017				
	(0.013)				
Week fixed effects	Yes	Yes	Yes	Yes	Yes
User fixed effects	Yes	Yes	Yes	Yes	Yes
Num. users	50,961	48,835	50,830	50,763	50,450
Num. obs.	5,962,437	5,713,695	5,947,110	5,939,271	5,902,650
Log Pseudo-likelihood	-6.730×10^6	-6.421×10^6	-6.725×10^6	-6.717×10^6	-6.083×10^6

Note. Column (1) includes user tenure as an additional control variable. Column (2) excludes nonparticipants who do not engage in any of the following activities during the post-announcement period: posting questions, voting for answers, subscribing to questions and topics, and following other users. This criterion results in the exclusion of 2,126 nonparticipants, which accounts for 4.26% of all nonparticipants. Columns (3) and (4) exclude nonparticipants who do not contribute any new answers after two/ four weeks of joining Zhihu. This results in the removal of 131 and 198 nonparticipants, respectively. Column (5) reports the results when we exclude highly active users who contribute fewer than 1,181 answers during our study period (i.e., the 99th percentile value). This criterion leads to the exclusion of 13 participants and 498 nonparticipants. Robust standard errors clustered at the user level are in the parentheses. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

B.3.2 Removal of Inactive Nonparticipants

Another potential concern is that the positive effect of the Zhihu Live program is driven by the inclusion of nonparticipants who are completely inactive or exit the platform soon after their first contributions. To address this concern, we use two criteria to identify such nonparticipants. First, we identify nonparticipants who do not engage in any major activities – posting questions, voting on answers, subscribing to questions/topics, and following users – during the post-period. Second, we identify nonparticipants who do not contribute any new answers after the first two or four weeks of joining Zhihu. We find that the identified churned users under each criterion account for only less than 5% of nonparticipants, and excluding them from the control group has little impact on our estimation results. Details of the results are in Column (2) – (4) of Table W6.

B.3.3 Exclusion of Outliers

The highly skewed distribution of our dependent variable, $Answer_{it}$, implies a possibility that some users may write more frequently than others, and, thus, have a disproportionately large influence on our estimation results. Therefore, we repeat our main analysis after excluding users who contributed more than 1,181 answers during our observation period, which corresponds to the 99th percentile value. This rule results in the exclusion of 13 participants and 498 nonparticipants, which account for 1.12% of participants and 1.00% of nonparticipants, respectively. We find that our main results are not sensitive to the exclusion of highly active users. Details of the estimation results are in the last column of Table W6.

B.3.4 Additional Functional Forms

Our main specification in Equation (1) uses Poisson regression to estimate the program effect, as our dependent variable has a skewed distribution (Athey and Imbens 2006). Another commonly used alternative for estimating such cases is to first log-transform the count variable(s) and then estimate the resulting model via a standard OLS regression (Angrist and Pischke 2009). Though Silva and Tenreyro (2006) show that such estimators are known to provide biased estimates of the true treatment effect, we still test the robustness of our Poisson regression estimates to log-linearized regressions owing to their high prevalence in the literature. We find that the results, as shown in Table W7, are directionally consistent with our earlier findings, though with notable differences in the magnitudes of the effect sizes.

Table W7: Average Treatment Effect of Zhihu Live Program on Answer Contributions: Robustness to Other Functional Forms

	Program announcement as the breakpoint		Participant actual enrollment as breakpoints	
DV:	$Ln(Answer_{it}+1)$	$Answer_{it}$	$Ln(Answer_{it}+1)$	$Answer_{it}$
$EP_i \times After_t$	0.171***	0.508***		
	(0.011)	(0.064)		
$EP_i \times After_{it}$			0.153***	0.457***
			(0.012)	(0.063)
Week fixed effects	Yes	Yes	Yes	Yes
User fixed effects	Yes	Yes	Yes	Yes
Num. users	50,961	50,961	50,961	50,961
Num. obs.	5,962,437	5,962,437	5,962,437	5,962,437
R^2	0.034	0.012	0.034	0.012
Model	log-linear regression	linear regression	log-linear regression	linear regression

Note. Robust standard errors clustered at the user level are in the parentheses. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

C Mechanism

C.1 Extract Question and Answer Characteristics

In this section, we investigate the characteristics of questions. We are mainly interested in two features: (1) the topical category that each question (and its answers) belongs to and (2) whether a question is open- or closed-ended.

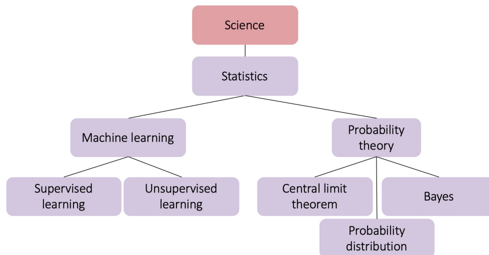
C.1.1 Classify Questions and Answers by Topical Category

Following the approach in Liu and Cong (2022), we classify topic tags into the 17 topical categories featured in the Zhihu Live market, including Internet, Psychology, Education, Travel, Lifestyle, Food & Cuisine, Career, Reading & Writing, Music & Games & Movies, Law, Business, Economics & Finance, Healthcare, Architecture & Interior Design, Art Appreciation, Science & Technology, and Sports. This classification allows us to categorize each question and its answers into the 17 categories according to the category affiliation of tags linked to each question. Furthermore, this approach enables us to determine the topical alignment between an answer and a specific Live event.

Our classification leverages existing relationships between tags on Zhihu's "topic tree," a hierarchical structure where each tag is related to certain parent tag(s) and child tag(s). Figure W7 illustrates a simplified example. All of the 17 categories correspond to certain parent tags on the tree, but some correspond to a single parent tag (e.g., the "Law" category corresponds to the parent tag "Law") and some consist of multiple tags (e.g., the "Science & Technology" category is combined with two parent tags, "Science" and "Technology"). Therefore, we define a many-to-one mapping, which stores every category and its corresponding parent tag(s) on the tree. This means that child tags under different parent tags can be classified into the same category.

Our classification contains the following steps. First, we adopt the Breadth First Search (BFS) algorithm to traverse every tag's parent tags on the tree, until we find the first parent tag that is in the tag-category mapping. By the nature of BFS, the first parent tag found has the shortest distance to the target tag on the topic tree. Therefore, we classify the target tag to the same category corresponding to the first parent tag found. A total of 28,763 tags are classified in this step.

Figure W7: An Example of Zhihu Topic Tree



Second, we use a Graph Convolutional Network (GCN) to classify the remaining tags. The GCN is designed to effectively extract node features from graph-structured data (Kipf and Welling 2016). In our application, the GCN learns meaningful representations for each tag by considering both its own features and the features of its neighboring tags. We initiate the features of nodes using pre-trained Zhihu tag embeddings from the Chinese-Word-Vectors repository developed by Li et al. (2018). This repository includes embeddings for 22,612 tags in our data.[2] For tags not existing in the repository, we initiate them with the average embedding of the 22,612 existing tags. We train the GCN model using a Deep Graph Infomax (DGI) (Veličković et al. 2018), a popular unsupervised learning algorithm on graphs. We use the trained GCN model to generate 128-dimensional embeddings for all tags. Then, for every uncategorized tag in the first step, we calculate its cosine similarity with every parent tag in the tag-category mapping and find the parent tag with the highest similarity. The tag is then classified into the same category as this parent tag.[3]

Finally, we classify questions according to the categorization of their tags. Specifically, each

[2] The Chinese-Word-Vectors repository offers an extensive collection of Chinese word embeddings that are trained on texts from various domains such as Sine Weibo, Baidu Encyclopedia, Chinese Wikipedia, and Zhihu. Its Zhihu word embeddings are trained on a large dataset consisting of 32,137 answers and 3,239,114 questions from Zhihu. These embeddings have demonstrated strong performance in sentiment classification and other downstream tasks (Qiu et al. 2018).

[3] Though every tag has a word embedding generated after the GCNs training, tags with no edges on the topic tree and no pre-trained embeddings only gain nonsense embeddings. We classify these tags as the "Others" category. In addition, tags that do not have pre-trained embeddings and only connect to the "Uncategorized" tag in the topic tree are also classified as "Others." A total of 23,181 topics are classified as "Others." For our study purpose, we exclude the "Others" category and only consider the 17 categories featured in the Zhihu Live market.

question is classified into the category that its topic tags belong to. A question may belong to more than one category if its topic tags are associated with multiple categories. Answers to a particular question are assigned to the same category as the question.

Results. We used the procedure described above to classify all of the 12,924,987 questions created on Zhihu during our study period. The classification result shows that questions are widely distributed among categories, with Internet, Music & Games & Movies, Healthcare, and Science & Technology having relatively larger proportions.

C.1.2 Identify Open-ended Questions

We use a supervised learning algorithm to classify questions as open- or closed-ended.

Table W8: Examples of Different Types of Questions in DuReader.

	Example1	Example2
Entity	On which day will iPhone be released?	What are the top10 movies in 2017?
Description	Why are firetrucks red?	How is Toyota Carola?
YesNo	Is 39.5 a high fever?	Does learning to play go improve intelligence?

Note. These examples are obtained from He et al. (2017).

Training Set Construction. Our training set is obtained from DuReader, a large-scale Chinese machine reading comprehension dataset (He et al. 2017). This dataset includes 200,366 questions from Baidu Zhidao, a large Q&A community similar to Zhihu. These questions are manually classified into one of three types: YesNo, Entity, and Description. Table W8 provides examples of the three types. The annotation of questions is completed by 52 experts and about 800 workers in a total of 51,408 man-hours. DuReader is an ideal source for constructing our training set, given the similarity between Baidu Zhidao and Zhihu questions and the comparability of its annotation design with our research objectives.[4]

Training Step. We use the entire set of 200,366 labeled questions from the DuReader database for our training process. For text analysis, we convert each question into a vector based on the

[4]An alternative method of building the training set involves randomly sampling questions from our data and labeling them ourselves. However, this approach is impractical due to the considerable effort and expertise it would require.

frequency of each word it contains. We accomplish this using the CountVectorizer from scikit-learn library in Python. This transformation produces a matrix where each row represents a unique question and each column represents a unique word. The value of each cell is the count of the corresponding word in that particular question.

We use XGBClassifier to learn the relationships between the extracted question features and labels. We use 80% of the questions as our training set. Given the imbalanced distribution of question types in the DuReader corpus, we implement random oversampling of the minority classes during training, using the RandomOverSampler function from the imbalanced-learn library in Python. The remaining 20% of the questions serve as a holdout sample to evaluate the performance of the trained model. The model achieves an accuracy of 85.7% in predicting question labels.

Prediction Step. We use the trained model to classify questions on Zhihu. To classify an unlabeled question, the classifier takes the vectorized question as input and outputs the predicted label: Entity, Description, and YesNo. For our study, we categorize questions as closed-ended if they are labeled as YesNo and open-ended if labeled otherwise.

Results. We classify all the 12,924,987 questions created on Zhihu during our study period. Of them, 2,612,526 are classified as close-ended and the rest are classified as open-ended.

Lastly, having identified both the question category and style features, we compare the questions in the hard and soft categories on several key characteristics. These include the weekly number of new questions, the weekly number of new open-ended questions, and the average number of answers per question. Detailed results are in Table W9.

Table W9: Comparing Questions in the Soft and Hard Category.

	Hard Category		Soft Category		t-Test, p-value
Variables	Mean	SD	Mean	SD	
Weekly num. of new questions	66,959	30,495	65,719	35,571	0.375
Weekly num. of new open-ended questions	52,790	16,048	53,854	29,608	0.435
Avg. num. of answers per question	2.16	0.48	2.61	0.52	0.000

Note. The hard and soft categories consist of 7,834,253 and 7,689,105 questions posted during our study period, respectively. The average number of answers per question is calculated based on answers received within the first week of question creation.

C.2 An Example of A Post-event Answer

Figure W8: An Example of A Post-event Answer

阅读 学习 教育 调查类问题 知乎 Live

参加知乎 Live 是怎样一种体验?

What is your experience with Zhihu Live?

已关注 写回答 邀请回答 14 条评论 分享 ···

137 个回答 137 answers 默认排序

匿名用户

293 人赞同了该回答

I streamed my fourth Live event yesterday. 25,000 listeners attended the live-streaming.

昨天完成了一场两万五千人的live，虽然是我的第四场live了，但是昨天比第一场还要紧张，手里拿着逐字稿，还是出现一些结巴的地方。昨天在live结尾处发了一个关于live的问卷调查，我勇敢地把我目前收到的反馈（包括建议和鼓励）分享一下。

题目\选项	1完全没用	2没用	3没感觉	4有帮助	5很有帮助	平均分
感受	1(1.1%)	1(1.1%)	4(4.4%)	59(64.84%)	26(28.57%)	4.19

虽然有两万五千人参与，但是问卷只收到93份，我自己个人感觉在live的过程中有大约200-300人在听，也就是其实现场在听的人数只有报名人数的1/100。

我回收的问卷就只有三道题，1.总体打分；2. 如果觉得不好，给出建议；3. 如果觉得有收获，收获了什么？

不得不说，我觉得我的听众素质都很高，即便是觉得没有收获的，但是还是很耐心地给我提出了建议，并没有恶语侮辱的情况出现，我心里还是很高兴的。

展开阅读全文

Click to read more

赞同 293 70 条评论 分享 收藏 喜欢 ···

FantasticCathy

首次公开募股（IPO）话题下的优秀答主

+ 关注

韦昌明、周源等 159 人赞同了该回答

I hosted one Live event last Friday at 9:00 pm. The topic is investment banking.

刚刚过去的周五（6月3日）晚上9:00，我做了一场Live，题目叫"投行那些事儿"。

问题征集帖是六天前发出去的，结果同志们纷纷在下面求链接...呃，其实，好像，貌似，理论上，求入口要在正式的预告贴下面，才算吧？

我的正式预告贴发的比较晚，周四下午才放出去，那时候离Live开始只有三十多个小时了，而且周四晚上有

@magasa

大大、

@傅渥成

先生的两场，周五早晨有

展开阅读全文

Click to read more

赞同 159 38 条评论 分享 收藏 喜欢 ···

C.3 Temporal Distance Effects across Users

Table W10: Answer Contributions over the Temporal Distance to Live Event: Heterogeneity across Users

	$Follower^1$	$Follower^2$	$Follower^3$	$Follower^4$	Nonrecipient	Recipient	Self	Invited	Soft-category	Hard-category
					Panel A: $DV = Answer_{it}$					
$AbsDistance_{it}$	-0.205^{***}	-0.137^{***}	-0.045^{**}	-0.002	-0.112^{***}	0.002	-0.102^{***}	-0.018	-0.096^{***}	-0.065^{***}
	(0.027)	(0.022)	(0.018)	(0.010)	(0.015)	(0.012)	(0.013)	(0.014)	(0.015)	(0.016)
Num. users	318	276	302	297	938	255	943	250	766	427
Num. obs.	3,740	3,432	5,035	5,667	12,586	5,288	12,384	5,490	10,938	6,936
Log Pseudo-likelihood	-5,856	-5,034	-6,342	-6,571	-17,916	-6,233	-17,985	-6,209	-15,026	-9,142
					Panel B: $DV = Upvote_{it}$					
$AbsDistance_{it}$	-0.063	-0.019	-0.049	-0.010	-0.022	-0.020	-0.039^{*}	-0.012	-0.061^{***}	0.013
	(0.063)	(0.058)	(0.030)	(0.025)	(0.023)	(0.024)	(0.020)	(0.030)	(0.020)	(0.029)
Num. users	222	214	267	259	734	228	751	211	615	347
Num. obs.	1,422	1,529	2,409	2,645	5,527	2,478	5,729	2,276	4,744	3,261
Log Pseudo-likelihood	-55,945	-53,934	-149,317	-319,768	-365,153	-232,410	-354,149	-250,680	-335,741	-259,909
					Panel C: $DV = TopicOverlap_{it}$					
$AbsDistance_{it}$	-0.047^{***}	-0.012	-0.029^{**}	-0.002	-0.028^{***}	-0.001	-0.024^{***}	0.014	-0.019^{*}	-0.005
	(0.017)	(0.014)	(0.014)	(0.012)	(0.009)	(0.010)	(0.010)	(0.010)	(0.011)	(0.013)
$Answer_{it}$	0.108^{***}	0.063^{***}	0.100^{***}	0.101^{***}	0.082^{***}	0.101^{***}	0.080^{***}	0.114^{***}	0.078^{***}	0.102^{***}
	(0.010)	(0.014)	(0.009)	(0.016)	(0.015)	(0.011)	(0.015)	(0.010)	(0.016)	(0.011)
Num. users	218	197	238	233	675	211	695	191	558	328
Num. obs.	1,400	1,475	2,320	2,551	5,322	2,424	5,535	2,211	4,549	3,197
Log Pseudo-likelihood	-1,518	-1,684	-2,555	-2,768	-5,891	-2,780	-6,132	-2,524	-5,124	-3,547

Note. Segment g ($g \in 1,2,3,4$) consists of participants whose value for $Follower_i$ is in the gth quantile of $Follower_i$. $TopicOverlap_{it}$ and $Upvote_{it}$ are set to a missing value if a user contributes zero answers in a given week. Because of the inclusion of individual fixed effects, the Poisson estimation drops users with only one observation or all zero outcomes in the temporal distance window. Therefore, sample sizes are slightly different across dependent variables. Robust standard errors clustered at the user level are in the parentheses. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

REFERENCES

Angrist JD, Pischke JS (2009) *Mostly harmless econometrics: An empiricist's companion* (Princeton university press).

Athey S, Imbens GW (2006) Identification and inference in nonlinear difference-in-differences models. *Econometrica* 74(2):431–497.

DiPrete TA, Gangl M (2004) 7. assessing bias in the estimation of causal effects: Rosenbaum bounds on matching estimators and instrumental variables estimation with imperfect instruments. *Sociological methodology* 34(1):271–310.

He W, Liu K, Liu J, Lyu Y, Zhao S, Xiao X, Liu Y, Wang Y, Wu H, She Q, et al. (2017) Dureader: a chinese machine reading comprehension dataset from real-world applications. *arXiv preprint arXiv:1711.05073* .

Kipf TN, Welling M (2016) Semi-supervised classification with graph convolutional networks. *arXiv preprint arXiv:1609.02907* .

Li S, Zhao Z, Hu R, Li W, Liu T, Du X (2018) Analogical reasoning on chinese morphological and semantic relations. *arXiv preprint arXiv:1805.06504* .

Liu J, Cong Z (2022) Express: The daily me versus the daily others: How do recommendation algorithms change user interests? evidence from a knowledge-sharing platform. *Journal of Marketing Research* 00222437221134237.

Manchanda P, Packard G, Pattabhiramaiah A (2015) Social dollars: The economic impact of customer participation in a firm-sponsored online customer community. *Marketing Science* 34(3):367–387.

Oestreicher-Singer G, Zalmanson L (2013) Content or community? a digital business strategy for content providers in the social age. *MIS quarterly* 591–616.

Qiu Y, Li H, Li S, Jiang Y, Hu R, Yang L (2018) Revisiting correlations between intrinsic and extrinsic evaluations of word embeddings. *Chinese Computational Linguistics and Natural Language Processing Based on Naturally Annotated Big Data: 17th China National Conference, CCL 2018, and 6th International Symposium, NLP-NABD 2018, Changsha, China, October 19–21, 2018, Proceedings 17*, 209–221 (Springer).

Rosenbaum P, Everitt B, Howell D (2005) Observational study. encyclopedia of statistics in behavioral science.

Rosenbaum PR (2002) Overt bias in observational studies. *Observational studies*, 71–104 (Springer).

Silva JS, Tenreyro S (2006) The log of gravity. *The Review of Economics and statistics* 88(4):641–658.

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

Veličković P, Fedus W, Hamilton WL, Liò P, Bengio Y, Hjelm RD (2018) Deep graph infomax. *arXiv preprint arXiv:1809.10341* .

Zhang S, Lee D, Singh PV, Srinivasan K (2022) What makes a good image? airbnb demand analytics leveraging interpretable image features. *Management Science* 68(8):5644–5666.