

**Does Identity Disclosure Help or Hurt User Content Generation?  
Social Presence, Inhibition, and Displacement Effects  
(Online Appendix)**

**Jingchuan Pu**

Smeal College of Business  
Pennsylvania State University, University Park, Pennsylvania 16802  
E-mail: [pjchuan@gmail.com](mailto:pjchuan@gmail.com)

**Yuan Chen\***

\*Corresponding author  
School of Information Management and Engineering  
Shanghai University of Finance and Economics, Shanghai 200433, China  
E-mail: [chen.yuan@mail.shufe.edu.cn](mailto:chen.yuan@mail.shufe.edu.cn)

**Liangfei Qiu**

Warrington College of Business  
University of Florida, Gainesville, Florida 32611  
E-mail: [liangfeiqiu@ufl.edu](mailto:liangfeiqiu@ufl.edu)

**Hsing Kenneth Cheng**

Warrington College of Business  
University of Florida, Gainesville, Florida 32611  
E-mail: [hkcheng@ufl.edu](mailto:hkcheng@ufl.edu)

## **Online Appendix A. Effects of Identity Disclosure on Users' Cases and Questions Generation**

In our research context, users can generate cases and questions, or review the existing cases and answer the existing questions. Because users mainly produce reviews and answers on the community in our research setting, we are particularly interested in the effects of identity disclosure on user content generation in terms of review and answer activities in our main analyses. In other words, we focus on users' responses to the existing cases and questions. The decisions on whether to provide a response and how much effort exerted on this response largely depend on the potential image received and the attributes of existing cases and answers. Therefore, by controlling for the number of and the topics of the existing cases and answers, our analyses can reveal how identity disclosure affects users' review and answer activities through its impact on image motivation.

In Online Appendix A, we check another interesting aspect of user content generation: the effects of identity disclosure on users' willingness to generate cases and questions. Different from users' review and answer activities which are largely driven by the perceived image and the available cases and questions, users' case and question generation can also be motivated by their need to share experience or seek help on the community. For instance, the attributes of the company's products may affect this need. In our context, identity disclosure was naturally implemented in the focal section and we analyze its effects by comparing users' activities before and after this change. We can also use our main models to identify the effects of identity disclosure on users' willingness to generate cases and questions. It is reasonable to assume that users' need to share experience (generate cases) or seek help (ask questions) is stable since the available products of the company in our study are not frequently changing. Nonetheless, without the control for the attributes of company's products, we need to interpret the effects of identity disclosure on users' cases or questions generation cautiously.

The outcomes we are interested here are the number of cases and questions generated by our sampled users. We generate two dependent variables,  $CaseCount_{it}$  and  $QuesCount_{it}$ , to respectively quantify the number of cases and the number of questions user  $i$  generates in month  $t$ . We estimate the effects of identity disclosure on users' cases and questions generation with the following models, which is similar to our main models described in equation (1):

$$\log(Y_{it} + 1) = \beta_0 + \beta_1 Year2014_t + UserExpCQ_{it} + \sum_{m=2}^{12} \gamma_m MonthDummy_m + \delta_i + \epsilon_{it} \quad (A.1)$$

where  $i$  and  $t$  indicate a user and a month respectively.  $Y_{it}$  is the aforementioned dependent variables of user  $i$  at month  $t$ . We take natural logarithms of the dependent variables plus one before including them in the estimation. The variable  $Year2014_t$  is the covariate indicating whether users' identities are disclosed in the focal section, and it equals 1 if the observation is for users' activities in 2014, and 0 otherwise. We include the variable  $UserExpCQ_{it}$ , which is the total number of cases plus total number of questions user  $i$  posted up until month  $t$ , to quantify the experience that user  $i$  has in terms of sharing cases and posting questions. Consistent with our main models, we include calendar month dummies and the user-level fixed effects  $\delta_i$ . We cluster the error terms at the user level to account for autocorrelation over time. Last,  $\beta_0$  is an intercept and  $\epsilon_{it}$  is a mean-zero random error term.

Table A.1 describes the descriptions and summary statistics of our main variables. We focus on the same 591 sampled individuals in our main analysis. The average pieces of cases generated by users is 0.117 per month. The sampled users ask 0.685 questions per month on average. The results in Table A.2 indicate that after identity disclosure in the focal section, users are less likely to generate cases (Table A.2, Model 1). Although the Q&A section remains anonymous during our study period, users post more questions after their identities are disclosed in the focal section (Table A.2, Model 2). These findings are aligned with the studies on the peer effects in the UGC context (Zhao et al. 2018), users' case and question behaviors are likely to be positively related to their perception on others' participation level in the review and Q&A sections. As the review volume substantially decreases after identity disclosure in the focal section, users are less likely to generate cases because the potential number of responses decreases.

**Table A.1 Main Variables at User-Month Level**

Variable	Description	Mean	Std. Dev.	Min	Max	N
$Year2014_t$	A dummy variable which indicates whether the observation describe the time period within 2014: 1 means the month $t$ is in 2014, 0 otherwise.	0.5	0.5	0	1	14,184
$CaseCount_{it}$	Number of cases user $i$ produces at month $t$ .	0.117	0.514	0	21	14,184
$QuesCount_{it}$	Number of questions user $i$ posts at month $t$ .	0.685	5.142	0	124	14,184
$UserExpCQ_{it}$	The total number of cases and questions user $i$ posts produces by month $t$ .	12.770	31.922	0	926	14,184

**Table A.2 The Effects of Identity Disclosure on Users' Cases and Questions Generation**

Dependent var.	(1)	(2)
	$CaseCount_{it}$	$QuesCount_{it}$
$Year2014_t (\beta_1)$	-0.127*** (0.007)	0.086*** (0.013)
$UserExpCQ_{it}$	0.203*** (0.020)	0.046*** (0.011)
Month Dummies	YES	YES
User Fixed	YES	YES
R-squared	0.343	0.346
Observations	14,184	14,184

Note. Robust standard errors clustered by each user are in parentheses. The unit of analyses is at user-month level. \* $p < 0.1$ ; \*\* $p < 0.05$ ; \*\*\* $p < 0.01$ .

## References

Zhao K, Zhang B, Bai X (2018) Estimating Contextual Motivating Factors in Virtual Interorganizational Communities of Practice: Peer Effects and Organizational Influences. *Information Systems Research* 29(4):910-927.

## Online Appendix B. Additional Analyses

In Online Appendix B, we conduct several additional analyses to further verify our theorization. First, we analyze the moderation effect of potential needs to compete with the peers and earn image. Second, by utilizing the answer category that doesn't provide virtual point, we find that the displacement effect is partially motivated by users' pursuit for the volume-based image represented by the virtual points. Finally, we break users into three different categories based on whether they only generate reviews, only generate answers, or generate both types of content in 2013. The results indicate that even for the users who only generate content in the neighbor section before identity disclosure, displacement effect might influence their answer behaviors via peer effects.

### B.1 Competition Effects

In the hypothesis development, we propose that identity disclosure in the focal section changes users' average content length in that section by increasing social presence. With identity disclosure, users can better gain image from the effort they invest in each content and they better sense the competition for attention in the focal section. In this light, how identity disclosure changes users' review activities is also related to the perceived competition from the peers in the focal section. To better attract attention and gain image, users should further increase their effort intensity when the peers also exert high level of effort.

We test how competition moderates identity disclosure's effects on users' activities in the focal section by adding one interaction in our equation (1):  $Year2014_t \times LenReview\_Peers_t$ , where  $LenReview\_Peers_t$  is the log-transformed average length of all reviews, regardless of the generator, in month  $t$ . A higher value of  $LenReview\_Peers_t$  represents higher need to compete with the other content generators for attention. Table B.1 reports the results. The coefficient of the interaction term is significantly positive when the dependent variable is  $NbrReview_{it}$  (Table B.1, Model 1) or  $LenReview_{it}$  (Table B.1, Model 2). These results indicate that the competition with peers might strengthen the role of identity disclosure on the review length but negatively moderate the intensity of its inhibition effect. We also use the average length

of the existing reviews of the user  $i$ 's reviewed cases up to the point user  $i$ 's review is posted (i.e.,  $LenExistingReview_{it}$ ) to measure the necessity of the competition for attention when he or she posts the review, the result is quantitatively consistent (Table B.1, Model 3). The results further complement our main results on the effects of identity disclosure by showing that users change their content-generation activities in the focal section to earn better image.

**Table B.1. The Effects of Peers' Effort**

Dependent var.	(1)	(2)	(3)
	$NbrReview_{it}$	$LenReview_{it}$	$LenReview_{it}$
$Year2014_t$	-6.091*** (0.721)	-2.080*** (0.361)	-1.499*** (0.222)
$Year2014_t$ $\times LenReview\_Peers_t$	1.729*** (0.209)	0.628*** (0.103)	-
$Year2014_t$ $\times LenExistingReview_{it}$	-	-	0.455*** (0.063)
$UserExp_{it}$	-0.007 (0.016)	-0.015 (0.014)	-0.016 (0.014)
$CaseCount_t$	0.000*** (0.000)	-	-
Month Dummies	YES	YES	YES
User Fixed	YES	YES	YES
R-squared	0.285	0.539	0.551
Observations	14,184	4,017	3,979

Note. Robust standard errors clustered by each user are in parentheses. The unit of analyses is at user-month level. \* $p < 0.1$ ; \*\* $p < 0.05$ ; \*\*\* $p < 0.01$ .

## B.2 Power of Virtual Point

We postulate that one reason of the displacement effect of identity disclosure is users' pursuit of volume-based image. As virtual point is the same from generating one review and posting one answer, but the effort of generating one review increases (i.e., increased social presence), users would choose to generate more answers. However, the virtual point is only for the work-related (i.e., marketing strategy and products) answers, and users would not get any point from the leisure-related answers. If the displacement effect is partially motivated by users' pursuit for the virtual points, we should expect that users generate both work-related and leisure-related answers more after identity disclosure, but the increase in work-related answers is larger. To test whether this is the case, we first estimate equation (1) with the log-transformed number of leisure-related answers,  $NbrLeisAnswer_{it}$ . We then create a dependent variable  $LeisAnswerRatio_{it}$ , which is

the ratio of number of leisure-related answers to the total answers for user  $i$  in month  $t$ . We include one additional control variable,  $LeisQRatios_t$ , which quantifies the ratio of the leisure-related questions to the total number of questions in month  $t$ . The results are reported in Table B.2. Specifically, the coefficient of  $Year2014_t$  is significantly positive when the dependent variable is  $NbrLeisAnswer_{it}$  (Table B.2, Model 1) and significantly negative when the dependent variable is  $LeisAnswerRario_{it}$  (Table B.2, Model 2). These results indicate that although users generate more both leisure-related and work-related answers, they are motivated by the pursuit of volume-based image and disproportionately generate more work-related answers to earn virtual points.

**Table B.2. The Effects of Identity Disclosure on Users' Answers with/without Virtual Points**

Dependent var.	(1)	(2)
	$NbrLeisAnswer_{it}$	$LeisAnswerRario_{it}$
$Year2014_t (\beta_1)$	0.080*** (0.028)	-0.081*** (0.013)
$UserExp_{it}$	0.027*** (0.011)	0.009*** (0.003)
$QuesCount_t$	0.000 (0.000)	0.000*** (0.000)
$LeisQRatios_t$	0.663** (0.336)	0.304** (0.106)
Month Dummies	YES	YES
User Fixed	YES	YES
R-squared	0.226	0.186
Observations	11,904	6,352

Note. Robust standard errors clustered by each user are in parentheses. The unit of analyses is at user-month level. \* $p < 0.1$ ; \*\* $p < 0.05$ ; \*\*\* $p < 0.01$ .

### B.3 Different Intensity of Displacement Effect

We propose that identity disclosure in the focal section affects users' content-generation behaviors in the neighbor section due to the displacement effect: users tend to displace the low-effort content from the focal section to the neighbor section. However, our main analyses show that after taking two sections together, users' content volume increases and content effort exerted on each content decreases overall, implying that displacement effect may be amplified in the UGC context. We perform a difference-in-differences (DID) analysis by breaking the users into three different categories based on the intensity of displacement effect. The users who generate reviews in 2013 are considered more likely to be affected by the displacement

effect of identity disclosure, including those who only write reviews and generate both reviews and answers in 2013. The users who only generate answers in 2013 are likely to focus only on the Q&A section. They are less likely to be directly affected by the displacement effect, because they couldn't displace their inhibited activities from the review section to the Q&A section. We create three binary variables to indicate which group user  $i$  is in:  $OnlyReview_i$  equals 1 if user  $i$  only contributes in the review section in 2013, otherwise equals 0;  $Both_i$  equals 1 if user  $i$  generates both reviews and answers in 2013, otherwise equals 0;  $OnlyAnswer_i$  equals 1 if user  $i$  only generates answers in 2013, otherwise equals 0. We alter our equations (1) to (B.1) by replacing the time intervention term,  $Year2014_t$ , with three interactions:  $Year2014_t \times OnlyReview_i$ ,  $Year2014_t \times Both_i$  and  $Year2014_t \times OnlyAnswer_i$ . Table B.3 reports the regression results.

$$\log(Y_{it} + 1) = \beta_0 + \beta_1 Year2014_t \times OnlyReview_i + \beta_2 Year2014_t \times Both_i + \beta_3 Year2014_t \times OnlyAnswer_i + X_{it} + \sum_{m=2}^{12} \gamma_m MonthDummy_m + \delta_i + \epsilon_{it} \quad (B.1)$$

For the users in group  $OnlyAnswer_i$ , the variable  $LenReview_{it}$  is missing for them at the months in 2013 (these users did not generate reviews in 2013). Thus, when the dependent variable is  $LenReview_{it}$  (Table B.3, Model 4), the estimation for the coefficient of  $Year2014_t \times OnlyAnswer_i$  will be missing because this estimation represents the difference between users' review length in 2013 and their review length in 2014 but the review length is missing value for the  $OnlyAnswer_i$  users in 2013. In this light, the estimation for the coefficient of  $Year2014_t \times OnlyReview_i$  will be missing when the dependent variable is  $LenAnswer_{it}$  (Table B.3, Model 5). It is worth noting that even though the users who only contribute in the neighbor section are not directly affected by the displacement effect, which is brought by the identity disclosure in the focal section, they also generate more answers but expend less effort on each content (the coefficients of  $Year2014_t \times OnlyAnswer_i$  are significant in Models 2 and 5, Table B.3). The users who are not directly affected by the displacement effect may change their behaviors in the neighbor section after observing the other users' activities. That is, the users can be indirectly affected by the displacement effect through the peer effects of the content generation (Lee et al. 2015, Burtch et al. 2017). These results provide additional

explanations on why users' content volume increases and content effort exerted on each content decreases overall after identity disclosure.

**Table B.3. DID Estimations of the Effects of Identity Disclosure**

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent var.	<i>NbrReview<sub>it</sub></i>	<i>NbrAnswer<sub>it</sub></i>	<i>NbrContent<sub>it</sub></i>	<i>LenReview<sub>it</sub></i>	<i>LenAnswer<sub>it</sub></i>	<i>LenContent<sub>it</sub></i>
<i>Year2014<sub>t</sub></i>	-0.106	0.918***	1.142***	0.119*	-	0.138**
× <i>OnlyReview<sub>i</sub></i>	(0.080)	(0.214)	(0.201)	(0.064)	-	(0.054)
<i>Year2014<sub>t</sub></i>	-0.517***	0.622***	0.750***	0.100***	-0.155***	-0.054***
× <i>Both<sub>i</sub></i>	(0.072)	(0.115)	(0.105)	(0.023)	(0.026)	(0.020)
<i>Year2014<sub>t</sub></i>	0.065	0.323*	0.788***	-	-0.129**	-0.168**
× <i>OnlyAnswer<sub>i</sub></i>	(0.064)	(0.184)	(0.180)	-	(0.068)	(0.067)
<i>UserExp<sub>it</sub></i>	-0.006***	0.362***	0.355***	-0.028*	-0.103***	-0.078***
	(0.017)	(0.051)	(0.050)	(0.015)	(0.014)	(0.012)
<i>CaseCount<sub>t</sub></i>	-0.001***	-	-	-	-	-
	(0.000)					
<i>QuesCount<sub>t</sub></i>	-	0.001***	-	-	-	-
		(0.000)				
<i>PostCount<sub>t</sub></i>	-	-	0.000***	-	-	-
			(0.000)			
Month Dummies	YES	YES	YES	YES	YES	YES
User Fixed	YES	YES	YES	YES	YES	YES
R-squared	0.282	0.545	0.520	0.525	0.550	0.504
Observations	11,904	11,904	11,904	3,807	6,352	7,151

Note. Standard errors are in parentheses. The unit of analyses is at user-month level. \* $p < 0.1$ ; \*\* $p < 0.05$ ; \*\*\* $p < 0.01$ .

## References

- Burch G, Hong Y, Bapna R, Griskevicius V (2017) Stimulating online reviews by combining financial incentives and social norms. *Management Science* 64(5):2065-2082.
- Lee Y J, Hosanagar K, Tan Y (2015) Do I follow my friends or the crowd? Information cascades in online movie ratings. *Management Science* 61(9):2241-2258.