MIS Quarterly

RESEARCH ARTICLE

NURTURING ONLINE COMMUNITIES: AN EMPIRICAL INVESTIGATION

Sofia Bapna and Mary J. Benner
Carlson School of Management, University of Minnesota,
Minneapolis, MN 55455 U.S.A. {sbapna@umn.edu} {mbenner@umn.edu}

Liangfei Qiu
Warrington College of Business, University of Florida,
Gainesville, FL 32611 U.S.A. {liangfeiqiu@ufl.edu}

Appendix A

Additional Tables

Table A1. An Example from Each High-Level Category of Firm Post Type

Post Type	Post Content
Conveys Credibility (product and industry knowledge)	"There several ways a wine can be considered "green." Have you tried one? How did it taste?" (with a link to a blog article titled "How Can I Find "Green" Wines?)
Conveys Professional Organizing (professional structure)	"It's time for the newest round of team member introductions… meet our new Social Media Manager <name deleted for privacy>! We're so happy to have her as part of our team."
Conveys Organizational Achievements (firm milestone partnership of award)	"In less than 24 hours Fab is going to be honored with not one, not two, but THREE Webby awards! We're beyond excited and extremely grateful to be receiving such high honors."
Seeks Opinions	"Would you rather: Nautical or Bohemian? Choose 1." (with an image of the two prints shown side by side)
Conveys Monetary Incentives (promotions or offers)	"It's official - $0 shipping on EVERY offer from 1-5pm! What will you pick up?"

Table A2. Correlation Matrix

	Likes	Com-ments	Shares	UGC Count	Firm Post Count	Google Trends	Acqui-sition Merger	New Product Line Added	Finan-cing Secured
Comments	0.72								
Shares	0.84	0.63							
UGC Count	0.03	0.14	0.13						
Firm Post Count	0.13	0.17	0.02	0.05					
Google Trends	0.26	0.21	0.22	0.04	0.09				
Acquisition Merger	0.07	0.04	0.09	0.10	-0.02	0.06			
New Product Line Added	0.02	0.05	0.23	0.15	-0.02	0.06	-0.05		
Financing Secured	0.01	0.02	-0.03	0.03	0.17	0.01	-0.03	-0.03	
Award	0.00	-0.01	0.00	0.01	-0.11	0.01	-0.05	-0.04	-0.03

Table A3. Collinearity Diagnostics

	VIF	SQRT VIF	Tolerance	R-Squared
Likes	4.97	2.23	0.20	0.80
Comments	2.16	1.47	0.46	0.54
Shares	4.28	2.07	0.23	0.77
UGC Count	1.10	1.05	0.91	0.09
Firm Post Count	1.12	1.06	0.90	0.10
Google Trends	1.08	1.04	0.92	0.08
Acquisition Merger	1.03	1.02	0.97	0.03
New Product Line Added	1.20	1.09	0.84	0.16
Financing Secured	1.04	1.02	0.97	0.03
Award	1.02	1.01	0.98	0.02

Mean VIF 1.90

Table A4. How Many Firms Had Each Type of Post

Post Type	Number of Firms
Capability of Key Members	9
Association with Experts	11
Tips or Suggestions	15
Industry Information	12
Design Origin Information	12
Collection or Picks	13
Professional Process	8
Professional Structure	7
Firm Milestone Partnership or Award	10
Product Award or Media Mention	11
Opinion	15
Promotion or Offer	15
Contest	12

Table A5. The Share of Posts for Each Firm

Firm	Share of Posts
Firm 1	6.02%
Firm 2	8.30%
Firm 3	6.65%
Firm 4	5.16%
Firm 5	7.05%
Firm 6	7.61%
Firm 7	10.79%
Firm 8	2.66%
Firm 9	5.20%
Firm 10	8.56%
Firm 11	3.52%
Firm 12	10.78%
Firm 13	4.10%
Firm 14	10.70%
Firm 15	2.89%

Table A6. Comparison of Coefficients in the Subsample Analysis

	t-statistic	*p*-value
Tips or Suggestions	86.70	0.00
Industry Information	236.63	0.00
Design Origin Information	391.65	0.00
Collection or Picks	402.93	0.00
Firm Milestone Partnership or Award	193.03	0.00
Opinion	338.00	0.00
Promotion or Offer	451.33	0.00

Appendix B

Additional Robustness Checks

Sentiment Analysis for UGC

It is conceivable that individuals might be joining online communities in order to complain. If this is the case, then the sentiment of user generated posts is likely to be negative. To mitigate this concern, we conducted a sentiment analysis on user generated posts and found that the overall sentiment of user generated posts is positive, thus suggesting that individuals are not joining brands' online communities to complain.

The sentiment analysis of user generated posts was conducted using long-term short term memory (LSTM) networks, a deep learning technique. LSTMs are a type of network that "remembers" previous data and makes decisions based on that knowledge. These networks are especially relevant in sentiment analysis because each word in a sentence has meaning based on the surrounding words, that is, previous and upcoming words (Mousa and Schuller 2017). Our analyses generated a sentiment value for each user generated post. A post's sentiment value can range from minus one to plus one. If the sentiment value is greater than zero, the post's sentiment is positive, otherwise it is negative. The more positive the sentiment of the post, the greater the sentiment value. Table B1 shows the descriptive statistics of sentiment values of user generated posts. We find that the mean and median of sentiment values are 0.75 and 1, respectively. Additionally, 92.29% of user generated posts have a sentiment value larger than zero, suggesting that most user generated posts have positive sentiment, thus suggesting that individuals are not joining brands' online communities to complain.

Table B1. Descriptive Statistics of Sentiment Values of User Generated Post

	Number of User Generated Posts	Mean	Std. Dev.	Min	Max	Median
Sentiment	22658	0.75	0.41	-0.59	1	1

Arellano–Bond Tests for Autocorrelation

We test whether the dynamic panel model is correctly specified. The Arellano–Bond tests for autocorrelation are reported in Table B2, and indicate that there is no serial correlation in the first-differenced disturbances.[1] In addition, both the Sargan and Hansen tests do not reject the null hypothesis (p value > 0.87 for both tests).[2] These test results suggest that our model specifications are based on instruments that are valid (Mileva 2007; Roodman 2009). Further, our results show that the coefficient for the lagged *L.Log(Community Size)* variable is positive and significant. This indicates that *Community Size* in the previous period is a good predictor of current *Community Size*, and hence pertinent to our model.

Table B2. Arellano-Bond Test for AR(1) and AR(2) in First-Difference

Order	z	Prob > z
AR(1)	-1.92	0.055
AR(2)	-0.58	0.564

[1]The Arellano-Bond test for autocorrelation has a null hypothesis of no autocorrelation, and is applied to the differenced residuals. We marginally reject the hypothesis at AR (1) (p value = 0.055). The test for AR (1) process in first differences is usually expected to reject the null hypothesis (Mileva 2007) and the test for AR (2) in first differences is more important (Mileva 2007; Roodman 2009). We do not reject the null hypothesis at AR(2) (p value = 0.564), implying that the model is not misspecified.

[2]The null hypothesis for these tests is that the instruments as a group are exogenous.

Accessing the Quality of the PSM

We implemented the following tests to assess the quality of the PSM conducted (results of the regression model that uses the dataset generated by the PSM are shown in Table 8, Model 3).

First, we performed *t*-tests of equality of means before and after the matching to check whether observable characteristics are balanced across posts in the treatment and control groups. The left panel of Table B3 shows clear evidence of covariate imbalance between groups before matching. After matching, the differences of means are no longer statistically significant (right column of Table B3), suggesting that matching helped reduce the bias associated with differences in observable characteristics. In other words, the treated and control posts are more similar in terms of these characteristics. Figure B1 shows the standardized percentage bias for each covariate before and after matching, and indicates that the covariate imbalance has been reduced after matching. This is consistent with the results of the *t*-tests shown in Table B3.

Second, prior matching literature suggests that a common support or overlap condition is a critical assumption in matching (e.g., Heinrich et al. 2010). Checking the overlap or region of common support between treatment and control groups can be done through a visual inspection of the propensity score distributions for both the treatment and control groups (Heinrich et al. 2010). Figure B2 shows the propensity score histogram for the treatment and control groups, and reveals a clear overlapping of the distributions between the two groups, indicating that our data meet the common support assumption.

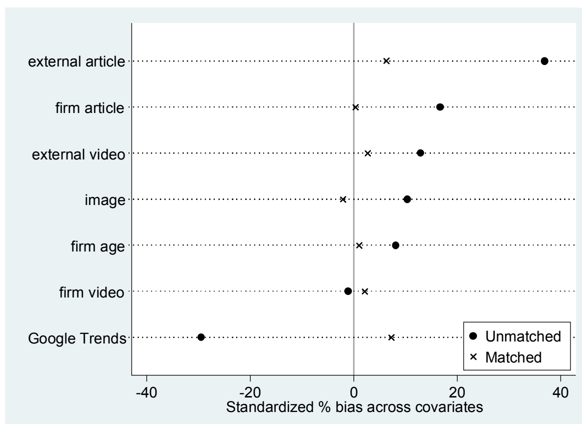

Figure B1. Standardized Percentage Bias for Each Covariate Before and After Matching

Table B3. Differences in Mean Before and After Matching

	Before Matching					After Matching				
	Mean Treated	**Mean Control**	**% Bias**	***t*-statistic**	***p*-value**	**Mean Treated**	**Mean Control**	**% Bias**	***t*-statistic**	***p*-value**
Firm Article	0.14	0.089	16.72	5.35***	0.00	0.14	0.14	0.32	0.07	0.95
External Article	0.17	0.058	36.93	13.53***	0.00	0.17	0.15	6.31	1.17	0.24
Firm Video	0.0021	0.0026	-1.11	-0.30	0.77	0.0021	0.0010	2.24	0.58	0.56
External Video	0.022	0.0066	12.92	4.99***	0.00	0.022	0.019	2.61	0.49	0.63
Image	0.87	0.83	10.44	2.94***	0.00	0.87	0.88	-2.02	-0.48	0.63
Firm Age	76.11	74.04	8.12	2.33**	0.02	76.11	74.86	1.03	0.23	0.82
Google Trends	31.32	38.86	-29.11	-8.45	0.00	31.32	29.45	7.31	1.67*	0.096

Note: ***p < 0.01, **p < 0.05, *p < 0.1

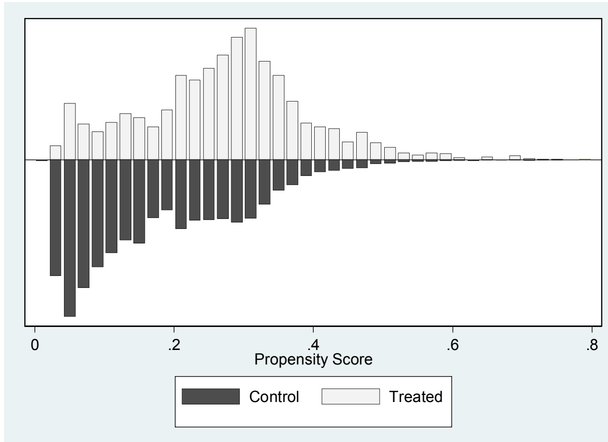

Figure B2. Propensity Score Histogram by Treatment Status

Rosenbaum Bounds: Sensitivity Analysis for PSM

We use PSM to try to address endogeneity concerns due to unobserved variables that might affect engagement as well as the explanatory variables. In our PSM analysis, we constructed control and treatment groups of matched posts and ensured that these groups were comparable on observable characteristics. However, it is possible that treatment and control groups differ on some unobserved characteristics (such as firm performance) even after matching on observed characteristics, leading to a hidden bias.

Rosenbaum's method of sensitivity analysis assesses if the estimates that are based on PSM are robust to the possible presence of an unobserved confounder. The sensitivity analysis provides a statement about the magnitude of the hidden bias that would need to be present to explain the associations actually observed (Rosenbaum 2002).

The sensitivity analysis considers several possible values of Γ, and shows how inferences might change based on the value of Γ. Γ is the odds ratio of treatment assignment, which is a measure of the degree of departure from a study that is free of hidden bias. Table B4 reports p-values from Wilcoxon signed rank tests for the averaged treatment effect while setting the level of hidden bias to different values of Γ.

When $\Gamma = 1$, we assume the absence of unobserved selection bias. In this case, both the upper and lower bounds of *p* values are zero (sig+ = 0, sig- = 0) indicating that the impact of post type is significant when there is no hidden bias. When $\Gamma = 8$, the upper bound of *p* value is greater than 5% (sig+ = .22, sig- = 0). Thus, we infer that for the impact of post type to disappear, the odds of differential assignment (to treatment and control group) due to unobserved factors is about 8. Intuitively, this means that for the impact of post type to disappear, the unobserved confounder has to cause a post to be eight times as likely as another post to receive treatment (assuming the two posts have the same observable characteristics). A Γ value of 8 is very large (Guo and Fraser 2010; Keele 2010), and it is well above the threshold values used in prior studies for robust PSM (e.g., Keele 2010; Wei and Lin 2017).

It is worth noting that the Rosenbaum bounds are "worst-case" scenarios. An insignificant upper bound *p* value for $\Gamma = 8$ does not mean that there is no true effect of post type when $\Gamma = 8$. This result means that the confidence interval for the effect of post type would include zero if an unobserved variable causes the odds ratio of treatment assignment to differ between treatment and control groups by 8.

In summary, this analysis suggests that the PSM estimates in our study are robust to the presence of unobserved confounders.

Table B4. Rosenbaum Bounds for PSM: Range of Significant Levels for the Signed Rank Statistic

Γ	Sig+	Sig-
1	0	0
2	2.4e-12	0
3	1.6e-10	0
4	3.9e-04	0
5	0.0026	0
6	0.012	0
7	0.036	0
8	0.22	0

Γ is the odds of differential assignment due to unobserved factors.
Sig+ is the upper bound significance level.
Sig- is the lower bound significance level.

Placebo Test

We conducted a placebo test to check if our results could be driven entirely by chance. Following the technique implemented by Bertrand et al. (2004), we randomly assigned post types to all the posts in our sample and reestimated our post level regression equation (described in the subsection "Second Phae: Measures and Methods" of the article). We repeated this exercise 1,000 times. Since the randomly assigned post types are fake, a significant "effect" at the 5% level should be found at most 5% of the time (50 times). From the 1000 runs, we find that the fraction of simulations in which the null hypothesis is rejected is 4.2% of the time (42 times). These results suggest that our results are unlikely to be driven by random chance, and thus help reduce concerns regarding identification of the effects described.

References

Bertrand, M., Duflo, E., and Mullainathan, S. 2004. "How Much Should We Trust Differences-in-Differences Estimates?" *Quarterly Journal of Economics* (119:1), pp. 249-275.

Guo, S., and Fraser, M. 2010. *Propensity Score Analysis: Statistical Methods and Applications*, Los Angeles: Sage Publications.

Heinrich, C., Maffioli, A., and Vásquez, G. 2010. "A Primer for Applying Propensity-Score Matching: Impact-Evaluation Guidelines," Technical Note No. IDB-TN-161, Office of Strategic Planning and Development Effectiveness, Inter-American Development Bank.

Keele, L. 2010. "An Overview of Rbounds: An R Package for Rosenbaum Bounds Sensitivity Analysis with Matched Data," Working Paper, Georgetown University (http://www.personal.psu.edu/ljk20/rbounds%20vignette.pdf).

Mileva, E. 2007. "Using Arellano-Bond Dynamic Panel GMM Estimators in Stata," unpublished paper, Economics Department, Fordham University.

Mousa, A., and Schuller, B. 2017. "Contextual Bidirectional Long Short-term Memory Recurrent Neural Network Language Models: a Generative Approach to Sentiment Analysis," in *Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics* (1), pp. 1023-1032.

Roodman, D. 2009. "How to Do xtabond2: An Introduction to Difference and System GMM in Stata," *The Stata Journal* (9:1), pp. 86-136.

Rosenbaum, P. R. 2002. *Observational Studies* (2nd ed.), New York: Springer.

Wei, Z., and Lin, M. 2017. "Market Mechanisms in Online Peer-to-Peer Lending," *Management Science* (63:12), pp. 4236-4257.