MIS Quarterly

RESEARCH ARTICLE

Free Versus For-a-Fee: The Impact of a Paywall on the Pattern and Effectiveness of Word-of-Mouth via Social Media

Hyelim Oh
School of Computing, National University of Singapore, 13 Computing Drive,
Singapore 117417 SINGAPORE {hyelim.oh@nus.edu.sg}

Animesh Animesh and Alain Pinsonneault
Desautels Faculty of Management, McGill University, Montréal, Québec, H3A 1G5 CANADA
{animesh.animesh@mcgill.ca} {alain.pinsonneault@mcgill.ca}

Appendix A

Summary of Underlying Theoretical Mechanisms

A. Impact of Paywall on WOM Pattern

Alternative Mechanisms		Rationale/Arguments/ Assumptions	Connected Literature/Theory	Resulting (Possible) Effect	Support for/Against
A1. Based on WTP and Exposure Theory	A1a	Light user segments (who are likely to have low WTP) are more likely to reduce (or even discontinue in extreme cases) their consumption of NYT content after the paywall implementation.	Utility theory — WTP (Danaher 2002)	Juxtaposing A1a and A1b leads to long tail effect due to the disproportionate reduction of popular content consumption (as a results of reduction of content consumption by light users).	A1a. Supported (see the descriptive statistics in Table 11). A1b. Supported (see results from the postestimation of finite mixture model in Table 9) Since the resulting effects as well as both the assumptions (A1a and A1b) are supported, we suggest that there is support for this mechanism.
	A1b	Light user segments are more likely to consume popular articles whereas the heavy user segment is more likely to consume a mix of niche articles and popular content.	Exposure theory (McPhee 1963)		
A2. Strategic User Behavior (by users with **low** WTP)	A2a	**Light user segments** (who are likely to have low WTP) are more likely to reduce (or even discontinue in extreme cases) their consumption of NYT content after the paywall implementation.	Utility theory — WTP (Danaher 2002)	Juxtaposing A2a and A2b, users who have lower WTP and are forced to curtail their NYT consumption due to paywall are more likely to curtail their consumption of popular content on NYT. This leads to long tail effect due to disproportionate reduction of popular content consumption (as a result of strategic reduction of popular content consumed by light users).	A2a is supported (as it is the same as A1a). A2b was not empirically tested in this paper but can be argued theoretically. To the extent that A2b is true, it is possible that this mechanism may also be playing a role in creating the long tail effect after the paywall.
	A2b	Search cost for finding popular content is lower than the search cost for finding niche content.	Search Cost Theory (differential search cost for popular and niche content on alternative website)		

A3. Strategic User Behavior (by users with **high** WTP)	A3a	**A segment of heavy users** decide to curtail their consumption of NYT in order to avoid paying subscription fee.		Leads to long tail effect due to the disproportionate reduction of popular content consumption (as a result of strategic reduction of popular content consumption by heavy users). This mechanism can coexist with exposure mechanism as well as the strategic user behavior by low WTP users.	Similar to the mechanism based on the strategic behavior of light users, assumption A3b was not empirically tested. However**,** our finding that the light users who read mostly popular content are more likely to reduce their content consumption after the paywall (see Table 11) suggests that this mechanism would have weak impact, if any.
	A3b	Search cost for finding popular content is lower than the search cost for finding niche content.	Search Cost Theory (differential search cost for popular and niche content on alternative website)		
A4. User Resentment (systematically related to reading behavior)	A4a	Due to negative emotional response toward the paywall, a proportion of NYT readers may boycott the NYT.	Campbell (1999); Xia et al. (2004)	Leads to long tail effect due to disproportionate reduction of popular content consumption (as a result of dropping out of users who feel resentment).	This mechanism does not suggest that there will be a differential impact of paywall on the content consumption of light and heavy users. Though we do not test assumption A4b, however, we find that consistent with exposure based mechanism light users reduce their content consumption more than heavy users after paywall (see Table 10), which cannot be explained by this mechanism. So we infer that this mechanism may have relatively weak role, if any.
	A4b	Users who feel resentment are more likely to read popular than niche content	No theory (suggested by an anonymous reviewer)		
B. Impact of Paywall on WOM Effectiveness					
Alternative Mechanisms		**Rationale/Arguments/ Assumptions**	**Underlying Theory/Rationale**	**Resulting (Possible) Effect**	**Support for/Against**
B1. Bypass effect	B1a	NYT allows visitors who come from links on social media to bypass its paywall. This bypass effect (i.e., increase in NYT non-subscribers' likelihood to click on a NYT content available through social media as they attempt to maximize the number of articles that can be accessed without paying subscription fee) may be dominant in website traffic generation.	Related to the Design of the Paywall mechanism	The relative strength of the positive relationship between social media WOM and website traffic may **increase** after a paywall implementation.	Our study does not directly measure the bypass effect but given the results that show **decrease** in the strength of the positive relationship between social media WOM and website trafficafter a paywall implementation, we can assume that role of such mechanism, if any, is small.
B2. Virality effect	B2a	Given that popular articles are the content that has a greater demand from a larger audience (Zentner et al. 2013), we expect that a decrease in the proportion of WOM about popular content will in turn lower the average clicks per link shared through social WOM.	Content characteristics play a significant role in determining the virality of online content (Berger and Milkman 2012)	The relative strength of the positive relationship between social media WOM and website traffic may decrease after a paywall implementation.	Results show that the paywall may lead to a disproportionate decrease in the WOM about popular content. Results also show decrease in the strength of the positive relationship between social media WOM and website traffic after a paywall implementation.
B3**.** User Resentment	B3a	Due to negative emotional response toward the paywall, a proportion of NYT readers may boycott the NYT and systematically do not click on and/or share NYT link on social media.	No theory (suggested by an anonymous reviewer)	The relative strength of the positive relationship between social media WOM and website traffic may decrease after a paywall implementation. Though this "resentment" based mechanism may be an alternative to the "virality "based logic suggesting negative impact of paywall on WOM effectiveness, both the mechanisms may coexist.	Our paper does not measure user resentment and cannot isolate the relative impact of virality vs. resentment mechanism on the relationship between social media WOM and website traffic. However, we believe that such a segment of users is likely to be relatively small.

Appendix B

Online Survey Questionnaires

1. Do you share online news articles (URLs) on social media (e.g., Facebook, Twitter)?
2. How many online newspaper articles (e.g., *The New York Times*, *The Los Angeles Times*) do you READ per week on average?
3. Please specify an approximate number of online news articles you READ in a normal week?
4. How many online newspaper articles (URL links) do you SHARE on social media (e.g., Facebook, Twitter)?
5. Please specify an approximate number of news articles (URL links) you SHARE on Facebook in a normal week.
6. Please specify an approximate number of news articles (URL links) you SHARE on Twitter in a normal week.

Appendix C

Supplementary Information of Difference-in-Difference Setup

Table C1. Key Demographics of NYT and LTAT Website Visitors

		NYT	LAT
Household income	< $30,000	16.32%	16.09%
	$30,000 – $59,000	27.21%	26.97%
	$60,000 – $99,999	23.60%	25.50%
	$100,000 – $149,999	17.22%	21.20%
	> $150,000	15.65%	10.24%
Age	18–24	8.35%	8.62%
	25–34	12.54%	11.49%
	35–44	14.59%	18.82%
	45–54	15.62%	21.50%
	55+	49.51%	37.13%
Gender	Female	43.56%	34.18%
	Male	56.44%	65.82%

Appendix C2. Comparison of NYT and LAT News Event Coverage

Before Paywall Rollout: Top 200 Most-Shared News Articles				
	NYT		**LAT**	
	# of News Articles	**# of Sharing**	**# of News Articles**	**# of Sharing**
Overlapped news events†	116	105,904	119	11,439
Region-specific local news events	5	2,019	26	2,340
Before Paywall Rollout: Top 200 Most-Shared News Articles				
	NYT		**LAT**	
	# of News Articles	**# of Sharing**	**# of News Articles**	**# of Sharing**
Overlapped news events‡	89	50,567	109	4,765
Region-specific local news events	8	3,166	39	3,681

†Major news events covered across NYT and LAT in pre-paywall sample include Japan tsunami & nuclear accident, Libya rebels & military actions, Bahrain & Arab protests, Wisconsin battle on union, U.S. government budget debates, and U.S. pacific tsunami.
‡Major news events covered across NYT and LAT in post-paywall sample include U.S. government shutdown & budget deficit, Japan nuclear disaster, Libya rebels & military actions, and election campaigns.

Appendix D

Supplementary Information of Finite Mixture Model Estimation

Table D1. Model Fit for Alternative Numbers of Segments

Number of Latent Segments	LL	AIC	BIC	R^2
1	-78369.01	156744.03	156768.58	0.0072
2	-75868.29	151750.59	151810.21	0.6602
3	-75069.08	150160.16	150253.84	0.7887
4	-75030.70	150091.40	150219.14	0.7930

Note: In terms of the BIC criterion, the models with greater numbers of segments improve model performance. However, an interpretation with a two-segment model is more suitable for our hypothesis testing, and adding an additional segment marginally improves the model fit indices after the three-segment model. We therefore opt to report the model estimates for both the two- and three-segment models.

Table D2. Parameter Estimates of Finite Mixture Models (Three-Segment Model)

Dependent Variable: ln(*Average Rank of Content Shared*)			
	Segment 1	**Segment 2**	**Segment 3**
Intercept	7.548 (.139)	5.041 (.096)	2.845 (.186)
ln(*User rank*$_j$)	-.624 (0.15)	.004 (.001)	.478 (.020)
Proportion	46%	51%	2%

Note: Standard errors in parentheses. *p < 0.1, **p < 005, ***p < 0.01.

Table D3. Postestimation of Finite Mixture Model (Three-Segment Model)

	Segment 1		Segment 2		Segment 3	
	Mean	**[95% conf. interval]**	**Mean**	**[95% conf. interval]**	**Mean**	**[95% conf. interval]**
ln(*Avg. Rank of Content Shared*$_j$)	1.782 (.008)	[1.765, 1.799]	5.161 (.006)	[5.148, 5.174]	7.376 (.007)	[7.360 ,7.391]
ln(*User rank*$_j$)	9.064 (.005)	[9.053, 9.075]	8.952 (.007)	[8.938, 8.966]	9.377 (.013)	[9.351, 9.402]

Note. Standard errors are in parentheses.

Appendix E

Supplementary Information on Robustness Checks

We estimate the following VARX model:

$$\begin{pmatrix} \ln(Website\,visit_t) \\ \ln(Tweets_t) \end{pmatrix} = \begin{pmatrix} \alpha_p \\ \alpha_w \end{pmatrix} + \begin{pmatrix} \varphi_{11} & \varphi_{12} \\ \varphi_{21} & \varphi_{22} \end{pmatrix} \begin{pmatrix} \ln(Website\,visits_{t-1}) \\ \ln(Tweets_{t-1}) \end{pmatrix} + \begin{pmatrix} \beta_p X_t \\ \beta_2 X_t \end{pmatrix} + \varepsilon_t$$

where ln(*Website visits*$_t$) denotes the daily gross site traffic at day t, and its one-day lagged variable is defined as *Website visits*$_{t-1}$. Similarly, ln(*Tweets*$_t$) represents the total number of tweets that contain the NYT link at day t. The dummy variables, *Saturday*$_t$ and *Sunday*$_t$, are included in all equations to control for variations due to differences in the type of news articles published on the weekends as well as the differences in the reading habits of consumers during the weekend. We first conducted the unit root tests. The Dickey-Fuller test results confirm that the variables are stationary rather than evolving in 95% confidence intervals. The results of the VARX model are reported in Table E1.

Table E1. Estimation Results for WOM and Website Traffic (VARX Model)

	Before Paywall	After Paywall
Site Traffic Equation: ln(*Website visits*$_t$)		
ln(*Tweets*$_{t-1}$)	0.2847 (0.1399*)	0.022 (0.0848)
ln(*Website visits*$_{t-1}$)	0.2283 (0.2648)	0.2343 (0.2025)
Saturday$_t$	-0.1648 (0.0548)	-0.1176 (0.037)***
Sunday$_t$	0.251 (0.0548)	-0.1176 (0.037)***
Constant	8.1302 (2.6188)***	10.6532 (2.4953)***
Tweets Equation: ln(*Tweets*$_t$)		
ln(*Website visits*$_{t-1}$)	0.7253 (0.5866)	-0.0430 (0.3063)
ln(*Tweets*$_{t-1}$)	0.1905 (0.3100)	0.1914 (0.1284)
Saturday$_t$	-0.4563 (0.1203)***	-0.3166 (0.0475)***
Sunday$_t$	-0.1504 (0.1214)	-0.4899 (0.0560)***
Constant	-1.9016 (5.800)***	10.7258 (3.7767)***

Standard errors are in parentheses. *p < 0.1, **p < 0.05, ***p < 0.01.