

A CROSS-CATEGORY ANALYSIS OF THE NETWORK EFFECTS ON WEBSITE USAGE

Yi-hsuan Chiang
Shih Hsin University
No. 1, Ln. 17, Sec. 1, Mucha Rd., Taipei 116, Taiwan
yihuan@cc.shu.edu.tw

Wen-jhan Jane
Shih Hsin University
No. 111, Sec.1, Mucha Rd., Taipei 116, Taiwan
wjane@cc.shu.edu.tw

ABSTRACT

This article addresses the issue of network effects derived from network externalities for website usage. A sample of 72,057 browsing behaviors was collected and examined by using Poisson models. Results show that total hourly pageviews of a website are positively related to an individual's hourly pageviews. Gender, age, income, and education are significant factors (e.g., males view more pages on the news, auction, searching and blog websites, and females view more pages on social networking, shopping, and video-sharing websites).

Keywords: Website Usage, Pageview, Browsing Behavior

1. INTRODUCTION

The success of online social networking service (SNS) websites, e.g., Facebook, Plurk, and Twitter, depends on the number and activity levels of their users. SNS users typically have numerous connections to other site members (i.e., "friends"), and those so-called friends may actually influence a member's website usage. Since individuals are connected through clicks, mutual influence often occurs. Scheinkman¹ defines "social interaction as particular forms of externalities, in which the actions of a reference group affect an individual's utility." In the information economics literature, the idea of a network effect is usually introduced to explain the interactions^{2, 3, 4}. It refers to demand of a good or service that increases with consumption of those products. Katz and Shapiro⁵ and Farrell and Saloner^{6, 7} indicate that a

consumer's utility increases as the number of consumers increases. Typical examples include telephone, fax, and intranet in a company.

One emerging setting for exploring the role of network effects in a competitive situation is in the adoption of online media for which compatibility and interoperability with others is crucial. However, insufficient research efforts were devoted to Internet economics in the early 2000's⁸. One of the major reasons may be inadequate data availability. The absence of the individuals' browsing behavior on the websites prevents us from analyzing many interesting phenomena on the Internet, e.g., information cascades, peer effects, network effects, etc. The network effect of Internet usage in our research focuses directly on the network externality. An externality is any situation in which the welfare of an individual is affected by the actions of other individuals, without a mutually agreed-upon compensation. For example, the benefit an individual gains from using a social networking site is related to the total number of people who use the site^{9, 10}. Goyal¹¹ and Jackson^{12, 13} offer exhaustive surveys of the growing literature on the economics of social networks. Our hope is that this paper will stimulate further empirical progress in this area.

In addition, gender is associated with information behavior and media use. For instance, women's total hours of television viewing is longer than men, but they watch fewer hours of news on television¹⁴. Previous studies also have found differences between women and men in news consumption through newspapers, news magazines, and television^{14, 15}. A survey study revealed gender differences in social media. The number of friends using SNSs was found to have a positive influence to continued usage among female users¹⁰. Another study used both qualitative and quantitative data that showed the effect of network externalities on corporate blog usage is greater for female than male employees¹⁶. Although a gender gap in access to the Internet has narrowed substantially in many countries, there still remain differences between women and men in certain activities.

The remainder of this paper is organized as follows. The next section describes the data and establishes the model for investigating network effects. Section three consists of the empirical results and brief discussion. The concluding remarks are summarized in the final section.

2. DATA DESCRIPTION AND METHODOLOGY

Internet usage is widespread in Taiwan. According to a recent report of the Taiwan Network Information Center (TWNIC)¹⁷, more than 17.6 million people (aged 12 and over) use the Internet in Taiwan. The Global Information Technology Report¹⁸ indicates that, for the 2013 Networked Readiness Index (NRI), Taiwan was ranked 10th, just behind the United

States. The position in world rankings reflects this country's preparedness and actual usage of information and communications technologies (ICTs), business and innovation environments, as well as the societal and economic impacts of ICTs. Taiwan is a representative environment for the study of network effects on Internet usage. Here, we use users' pageviews and their demographic characteristics to test network effects on browsing behaviors.

A dependent variable is usually assumed to be a continuous value in the regression. Unfortunately, our dependent variable, i.e., the user's hourly pageviews, is not. Poisson regression fits models of the number of counts/occurrences of an event. The basic Poisson distribution for i^{th} user's hourly pageviews (Y_i) can be written as

$$\text{Prob}(Y_i = y_i | x_i) = \frac{e^{-\lambda_i} \lambda_i^{y_i}}{y_i!} \quad (1)$$

where y_i is a particular count value. λ_i is the sole parameter representing the expected value of hourly pageviews, and $i = 1, \dots, N$ indexes the N users.

Our data includes observations for a group of users in a specific period. Therefore, we adjust the basic Poisson model and model the activity of i^{th} user's ($i = 1, \dots, N$) Internet usage over t ($t = 1, \dots, T$). The specification of the conditional hazard function is

$$\lambda_{it} = \exp(\mu_0 + X'_{it}\beta) \quad (2)$$

where λ_{it} represents the i^{th} user's Internet usage intensity in the t^{th} period. μ_0 is a constant and

$$X'_{it}\beta = \gamma_1 \text{NPView}_{i,t} + \gamma_2 C_i + \gamma_3 \text{DTime}_{i,t} + \varepsilon_{i,t} \quad (3)$$

C is a set of individual characteristics such as gender, age, education (Edu), and income. DTime represents the control variables of time "dummy." They include the dummy variables of the weekend (Saturday and Sunday = 1, otherwise = 0) and midnight (23:00-03:00 = 1, otherwise = 0). NPView is total pageviews, which is summed by the individual's hourly pageviews at time t on a site (PagevSite). γ_1 is used to measure the network effect on Internet usage.

Our data includes as many samples as possible for the purpose of representativeness in the analysis. The 20 largest websites in Taiwan are included, and they belong to seven categories. These websites are listed in Table 1 below.

Table 1. Websites in Taiwan included in the dataset and their functions

Website	Site Address	Category
Facebook	facebook.com	SNS
Plurk	plurk.com	SNS
Twitter	twitter.com	SNS
China Times	news.chinatimes.com	News
Apple Daily	nextmedia.com.tw	News
United Daily News	udn.com	News
Yahoo-Kimo News	tw.news.yahoo.com	News
MSN News	news.msn.com.tw	News
NOWnews	nownews.com	News
Yahoo-Kimo Shopping	buy.yahoo.com.tw	Shopping
PChome Shopping	shopping.pchome.com.tw	Shopping
PayEasy	payeasy.com.tw	Shopping
Yahoo-Kimo Bid	tw.bid.yahoo.com	Auction
Ruten	ruten.com.tw	Auction
Yahoo-Kimo Search	tw.search.yahoo.com	Searching
Google	google.com.tw	Searching
Bing	bing.com	Searching
Pixnet	pixnet.net	Blog
Wretch	wretch.cc	Blog
YouTube	youtube.com	Video Sharing

We used the longitudinal data consisting of pageviews and a sample users' profile across these websites in Taiwan from October 1-31, 2010. With InsightXplorer's Access Rating Online (ARO) measurement system and CyberPanel membership, 1088 users are traced across sites, and samples of research panel members' clickstream logs are recorded by a meter software. Therefore, by classifying the user's pageviews on different websites, a user's pageviews can be measured and matched by a total usage of a website. Because members do not remain on the same website all the time, the panel data in the period is unbalanced. In the sample, 57.37% of the individuals are female, and 7.2% have a master's degree. Total number in the sample is 72057.

Table 2 summarizes the descriptive measures of the data set. The average number of a user's hourly pageviews (PView) is 28.35, and a user's pageviews ranges from 1 to 981 pages. The average number of hourly pageviews of a website (PagevSite) is 4610.07, and total hourly pageviews of a website ranges from 1 to 18579 pages. This implies that, on average, approximately 163 ($= 4610.07/28.35$) respondents are on the website per

hour. The respondents are between the ages of 10 and 50, and the average age of respondents is 25 years old.

Table 2. Descriptive statistics (n = 72057)

Variable	Mean	Std. Dev.	Min	Max
PageView	28.35047	65.43062	1	981
PagevSite	4610.066	5528.345	1	18579
Gender	.4262736	.494538	0	1
Age	25.47039	7.917255	10	50
Education	3.800824	.6106298	1	5
Income	2.25824	1.851713	1	7
Weekend	.2342	.4235	0	1
Midnight	.2061285	.4045266	0	1

The regressions of the Poisson model proceeds in the following steps. First, total pageviews of a website (PagevSite) is simply included in the model for measuring the network effect. Then, the user's characteristic variables and time dummies are included in the regressions step by step to test robustness of the effect. Network effect theory posits that the adopter's demand derived from a network technology is positively associated with the size of the network^{19,20}. Therefore, the network effect (γ_1) is expected to be positively related to a user's pageviews.

3. EMPIRICAL RESULTS AND DISCUSSION

Individual's browsing behavior is regressed by equations (2) and (3). The full-sample results are presented in Table 3. The coefficients of total hourly pageviews of a website (PagevSite) are significantly positive. The significance of network effects shows a consistent effect on an individual's Internet usage. This indicates that the more other people use a website, the more an individual uses that same site. There exists a positive network effect on internet browsing, which matches previous expectations.

Table 3. The empirical results of network effects on Pageviews

Independent Variable	Individual's Website Pageviews				
	Model 1	Model 2	Model 3	Model 4	Model 5
PagevSite	.0001*** (0.0000001)	.0001*** (0.0000001)	.0001*** (0.0000001)	.0001*** (0.0000001)	.0001*** (0.0000001)
Gender		.222*** (.001)	.216*** (.001)	.222*** (.001)	.216*** (.001)
Age		.086*** (.0007)	.087*** (.0007)	.091*** (.0007)	.092*** (.0007)
SqAge		-.001*** (.00001)	-.001*** (.00001)	-.001*** (.00001)	-.001*** (.00001)
Income		.035*** (.0004)	.035*** (.0004)	.033*** (.0004)	.033*** (.0004)
Education		-.098*** (.001)	-.098*** (.001)	.	.
DEduJ				.065*** (.011)	.048*** (.011)
DEduH				-.134*** (.011)	-.141*** (.011)
DEduB				-.274*** (.011)	-.282*** (.011)
DEduM				-.264*** (.011)	-.273*** (.011)
Weekend			.116*** (.002)	.	.116*** (.002)
Midnight			-.047*** (.002)		-.046*** (.002)
Cons	2.626 *** (.001)	1.528*** (.010)	1.478*** (.010)	1.333*** (.014)	1.291*** (.014)
Log likelihood	-2029743.7	-1986706	-1983564.2	-1985559.4	-1982447.8
Pseudo R ²	0.184	0.201	0.203	0.202	0.203
Number of obs.	72057	72057	72057	72057	72057

Notes: 1. Standard error is in parenthesis.

2. *** indicates significance at 1%.

All the coefficients of characteristic variables are consistent from model 1 to model 5. The dummy variable of gender is significantly positive in all regressions, suggesting that males use these websites more than

females. For the coefficient of age, it is significantly positive related to a user's pageviews. The square term of age is significantly negative related to a user's pageviews. The results show that older people have more pageviews than younger people, but the effect increases at a decreasing rate. As for the results of income and education, the former is positively significant and the latter is negatively significant. This indicates that higher income (/less educated) people view more pages than lower income (/more educated) people. Specifically, in the analysis of education dummy, it shows that people who have a junior high school diploma view more pages. Moreover, the dummy variables of weekend and midnight are significantly positive and negative in all regressions, respectively. This suggests that browsing behavior on websites is higher on Saturday and Sunday, and it is lower at midnight.

These findings have important economic implications. Other people's behavior is an important factor influencing an individual's website pageviews. This result supports the hypothesis for network effects. There exists a network externality in the individual's usage of websites. The marginal effect which we estimated shows that an increase of 10000 external pageviews leads to 1.1 additional viewers. Eventually, the positive marginal effect makes the larger websites larger, and the smaller ones may be eventually driven out of the market. Therefore, the direction of regulation needs to focus on setting a level of competition in the market, and antitrust acts are necessary in order to keep a degree of diversification of websites.

After finding evidence of positive network effects on user's browsing behavior, we consider a further analysis of different types of websites. The data for seven types of websites are regressed respectively, and the comparisons of the influencing effects between them provide some interesting phenomena and policy implications.

As for the results of network effects on a user's pageviews in the SNS, news, shopping, auction, searching, blog, and video-sharing websites, all coefficients of PagevSite are positively significant, and they correspond with previous findings. The significance for both variables in these regressions reinforces the evidence of network effect on an individual's Internet usage.

Gender effects are different in the regressions for different types of websites. The coefficients of gender are significantly positive related to user's pageviews in the regressions of news, auction, searching, and blog websites, but they are significantly negative related to a user's pageviews in the regressions of SNS, shopping, and video-sharing websites. Males view more pages than females in the news, auction, searching, and Blog websites, but females view more pages than males in the SNS, shopping, and

video-sharing websites. That is, females prefer to view social networking, shopping, and video-sharing websites, and males prefer to view news, auction, searching, and blog websites.

The coefficients of education and income are inconsistent in the regressions for different types of websites. The coefficients of education are significantly positive related to a user's pageviews in the regressions of SNS and searching websites, but they are significantly negative related to a user's pageviews in the regressions of news, shopping, blog, and video-sharing websites. The coefficients of income are significantly positive related to a user's pageviews in the regressions of auction, searching, blog, and video-sharing websites, but they are significantly negative related to a user's pageviews in the regressions of SNS and news websites. More educated (/higher income) people view more pages than less educated (/lower income) people in the SNS and searching (/auction, searching, blog, and video-sharing) websites, but more educated (/higher income) people view less pages than less educated (/lower income) people in the news, shopping, Blog, and video-sharing (/SNS and news) websites. That is, more educated (/higher income) people prefer to view SNS and searching (/auction, searching, blog, and video-sharing) websites.

Other findings include the significance of age effect in the regressions of auction websites, and the significance of inverse age effect in the regressions of searching websites. The former evidence indicates that older people have more pageviews than younger people in auction websites, and the age effect increases at an increasing rate. The latter one shows that younger people have more pageviews than older people in searching websites, and the age effect decreases at an increasing rate. Moreover, the dummy variables of weekend are consistently positive and significant in the regressions, suggesting that browsing behavior on websites is higher on Saturday and Sunday.

4. CONCLUSIONS

With the limited empirical references available, establishing a methodology of investigation for network effects on individual's browsing behavior is a challenging task. Panel data of individuals' Internet usage and their characteristics on the 20 largest websites in Taiwan were collected, and the Poisson model was employed to investigate network effects. Results show empirical support for network effects in a website across different site categories. Furthermore, gender, age, income, and education are significant factors in the whole sample analysis. Male, older, higher-income, and less-educated users browse more pages on websites.

Other interesting findings are as follows: (1) Males view more pages on the news, auction, searching, and blog websites, and females view more pages on the SNS, shopping, and video-sharing websites. (2) Younger people have more pageviews than older people in searching websites, and the age effect decreases at an increasing rate. (3) More educated people prefer to view SNS and searching websites, and higher income people prefer to view auction, searching, blog, and video-sharing websites. (4) Browsing behavior on websites is higher on the weekend than during weekdays.

This research suggests some preliminary implications for understanding Internet usage from a network effects perspective. That is, an individual user's perceived benefit will increase as the quantity of other users adopting the website grows, and, accordingly, network externality influences pageview decisions. More total pageviews bring more individuals' visits to this site. This phenomenon is consistent with the communication technology adoption described by Strader et al.²¹.

Examining the dynamics of network effects is essential in developing business strategies and market reform, and predicting impacts on diversity of online media consumption. A website having more traffic will result in more individuals using the website. There exists a network externality on website usage that causes large websites to grow as smaller ones may be driven out of the market. For web businesses to get stronger is to merge with other companies. Therefore, government regulation needs to consider ensuring fair competition in the marketplace if the smaller creative websites need to be protected and enhanced.

Future research needs to explore how network effects contribute to adoption decisions. Like most previous research, we did not examine the boundaries of network effects. The effect has traditionally been viewed as linear. Research is needed to identify the threshold for effects, as well as what happens when an increasing number of users no longer create network effects. Moreover, understanding how potential adopters generate their perceptions of network participation benefits is crucial. Which groupings of user-types (e.g., education) generate benefits most in Internet usage? Do heavy users benefit more in their heavy usage? These important issues deserve further research.

REFERENCES

- [1] J.A. Scheinkman, Social interactions. In Steven N. Durlauf and Lawrence E. Blume (Eds.), *The new palgrave dictionary of economics* (online). Basingstoke: Palgrave Macmillan, 2008. <http://dx.doi.org/10.1057/9780230226203.1558>.
- [2] Z. Katona, P.P. Zubcsek, and M. Sarvary, Network effects and personal

- influences: The diffusion of an online social network. *Journal of Marketing Research*, 48(3), p425-443, 2011. <http://dx.doi.org/10.1509/jmkr.48.3.425>.
- [3] G.G. Parker, and M.W. Van Alstyne, Two-sided network effects: A theory of information product design. *Management Science*, 51(10), p1494-1504, 2005. <http://dx.doi.org/10.1287/mnsc.1050.0400>.
- [4] K. Zhu, K.L. Kraemer, V. Gurbaxani, and S.X. Xu, Migration to open-standard interorganizational systems: Network effects, switching costs, and path dependency. *MIS Quarterly*, 30(1), p515-539, 2006.
- [5] M.L. Katz, and C. Shapiro, Network externalities, competition, and compatibility. *American Economic Review*, 75(3), p24-440, 1985.
- [6] J. Farrell, and G. Saloner, Standardization, compatibility, and innovation. *The RAND Journal of Economics*, 16(1), p70-83, 1985. <http://dx.doi.org/10.2307/2555589>.
- [7] J. Farrell, and G. Saloner, Installed base and compatibility: Innovation, product preannouncements, and predation. *American Economics Review*, 76(5), p940-55, 1986.
- [8] M. Cave, and R. Mason, The economics of the Internet: Infrastructure and regulation. *Oxford Review of Economic Policy*, 17(2), p188-201, 2001. <http://dx.doi.org/10.1093/oxrep/17.2.188>.
- [9] DeAndrea, N.B. Ellison, R. LaRose, C. Steinfield, and A. Fiore, Serious social media: On the use of social media for improving students' adjustment to college. *The Internet and Higher Education*, 15(1), p15-23, 2012. <http://dx.doi.org/10.1016/j.iheduc.2011.05.009>.
- [10] K.Y. Lin and H.P. Lu, Why people use social networking sites: An empirical study integrating network externalities and motivation theory. *Computers in Human Behavior*, 27(3), p1152-1161, 2011. <http://dx.doi.org/10.1016/j.chb.2010.12.009>.
- [11] S. Goyal, *Connections: An introduction to the economics of networks*. Princeton: Princeton University Press, 2007.
- [12] M.O. Jackson, The economics of social networks, In Richard Blundell, Whitney Newey, and Torsten Persson (Eds.), *Advances in economics and econometrics, theory and applications, ninth world congress of the econometric society* (pp1-56). Cambridge: Cambridge University Press, 2006.
- [13] M.O. Jackson, *Social and economic networks*. Princeton: Princeton University Press, 2008.
- [14] H.W. Stanley, and R.G. Niemi, *Vital statistics on American politics, 1997-1998*. Washington, D.C.: Congressional Quarterly, 1998.
- [15] L.L. Bennett, and S.E. Bennett, Enduring gender differences in political interests: The impact of socialization and political dispositions. *American Politics Research*, 17(1), p105-122, 1989. <http://dx.doi.org/10.1177/1532673X8901700106>.

- [16] S. Wattal, P. Racherla, and M. Mandviwalla, Network externalities and technology use: A quantitative analysis of intraorganizational blogs. *Journal of Management Information Systems*, 27(1), p145-174, 2010. <http://dx.doi.org/10.2753/MIS0742-1222270107>.
- [17] TWNIC, *2014 Taiwan broadband usage survey*. Retrieved on March 1, 2015, from <http://www.twnic.net.tw/download/200307/20140820e.pdf>.
- [18] B. Bilbao-Osorio, S. Dutta, and B. Lanvin, *Global information technology report 2013. Growth and jobs in a hyperconnected world*. Geneva: World Economic Forum and INSEAD. Retrieved on August 10, 2014 from http://www3.weforum.org/docs/WEF_GITR_Report_2013.pdf.
- [19] M.L. Katz, and C. Shapiro, Technology adoption in the presence of network externalities. *Journal of Political Economy*, 94(4), p822-841, 1986. <http://dx.doi.org/10.1086/261409>.
- [20] J. Church, and N. Gandal, Complementary network externalities and technological adoption. *International Journal of Industrial Organization*, 11(2), p239-260, 1993. [http://dx.doi.org/10.1016/0167-7187\(93\)90006-X](http://dx.doi.org/10.1016/0167-7187(93)90006-X).
- [21] T.J. Strader, S.N. Ramaswami, and P.A. Houle, Perceived network externalities and communication technology acceptance. *European Journal of Information Systems*, 16, p54-65, 2007. <http://dx.doi.org/10.1057/palgrave.ejis.3000657>.

