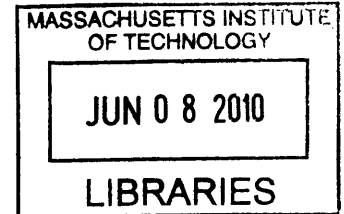# Newspapers and the Internet: Friends or Foes?
Evidence of Concentration of Choice in Korea

By

Jun Mo Park

Bachelor of Business Administration
Yonsei University, 2001

Submitted to the MIT Sloan School of Management
In Partial Fulfillment of the Requirements for the Degree of

ARCHIVES

Master of Business Administration

At the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2010

Signature of Author: _____

Jun Mo Park
MIT Sloan School of Management
May 7, 2010

Certified by: _____

Juanjuan Zhang
Assistant Professor of Marketing
Thesis Advisor

Accepted by: _____

Debbie Berechman
Executive Director, MBA Program
MIT Sloan School of Management

Page left intentionally blank

Newspapers and the Internet: Friends or Foes?
Evidence of Concentration of Choice in Korea

By

Jun Mo Park

Submitted to the MIT Sloan School of Management
On May 7, 2010 in Partial Fulfillment of the
Requirements for the Degree of
Master of Business Administration

Abstract

Does the expansion of the Internet provide opportunities for newspapers to attract more readers, or does it threaten their sustainability by cannibalizing subscription revenue from print circulation? To examine these competing hypotheses, I analyzed a unique data set of the monthly number of clicks on the homepages of the Web sites of four economic newspapers in Korea between 2001 and 2009. Using a unique quasi-experimental observation of the Korean media industry, I illustrated the differential impact of the Internet on leading newspapers and their followers. By attracting people to a common "playground" and providing inference information–namely information on others' decisions–the Internet has a significant impact on consumer choice of information products, reinforcing informational cascades through the inference effect. My findings reveal that since the advent of the Internet, customer preference has consistently been dominated by market leaders, strengthening the brand image of the leaders. This phenomenon will, over time, widen the gap between leading newspaper groups and laggards, providing new opportunities for market leaders only. Thus, my findings point to significant marketing strategies based on consumer choice.

Thesis Advisor: Juanjuan Zhang
Title: Assistant Professor of Marketing

Page left intentionally blank

1. Introduction

In October 2009, Rupert Murdoch, chairman of News Corporation and global media mogul, accused Google of stealing content from his news sites by linking to his news stories without compensating News Corp. In addition, Murdoch said the aggregator and plagiarists will soon have to pay a price for co-opting his content assets, urging media outlets to fight back against "stealing" their copyrighted material. Erik Schmidt, CEO of Google, sees the situation differently. According to Google, the search-engine giant sends greater than 300 million clicks a month to newspaper Web sites just by indexing their headlines, a line or two of content, and providing links to the story's Web site. Are newspapers and the Internet friends or foes?

In the Internet age, on-line news sites have become a major outlet for consumption of international and national news. According to a survey by the Pew Research Center, 40% of respondents obtained most of their news about national and international events from the Internet. The authors also found that consumers relied more on the Internet than on printed newspapers for their news. Looking at these trends, many industry observers have declared the "death of the newspaper." For example, former dean at the Columbia University Graduate School of Journalism, Tom Goldstein, suggested, "Unless they urgently respond to the changing environment, newspapers risk early extinction."

However, others believe that the Internet presents newspapers with an opportunity to compensate for lost circulation and readership. In fact, many major newspaper outlets are making use of the Internet to expand their newspaper's reach. Some U.S. examples are as follows:

- 92% of America's top 100 papers now offer news video on their Web sites.
- 95% of newspapers offer at least one reporter blog.
- One-third of newspapers now allow comments on their articles.
  (Source: http://www.bivingsreport.com/)

2. Literature Review

In general, prior research on the relationship between the Internet and newspapers argues that on-line newspapers are a problematic replacement for printed versions because they weaken subscription revenues. Several researchers have expressed concerns about on-line news cannibalization when readership habits are changing as consumers turn to the Internet for free news and information. Lapo Filistrucchi (2005) estimates a model of demand on market data from 1976 to 2001 for the five leading national newspapers in Italy and finds that opening a Web site has had a negative impact on market share. In research by Ulrich Kaiser (2005), on-line newspapers have had significant negative effect on the print circulation of national newspapers, but a significantly positive effect on local newspapers. A recent study by Thomas E. Patterson (2007) explains how the Internet poses a great threat to the newspaper industry through information consumption in American schools. He surveyed 1,262 ethics and sociology teachers on the way information media is used in American schools and found that Internet-based news sites are more often used as a mode of classroom instruction than daily newspapers. In addition, governmental policy and the U.S. Congress have considered possible federal action to bolster the shrinking print newspaper industry. For example, in a report prepared for Members and Committees of Congress, Suzanne M. Kirchhoff (2009) provides an excellent overview of the turmoil in the U.S. newspaper industry. This research underlines that policymakers may determine that some actions could ease the combination of technological shift (the Internet) and the recession-related financial problems in the newspaper industry.

In contrast, several studies have shown that the advent of on-line newspapers does not appear to threaten the survival of print media. Matthew Gentzkow (2007) presents an empirical framework to explore the relationship between the print and on-line newspaper readership on consumers in Washington, DC. He suggests that even though print and on-line papers are obviously substitutes and the magnitude of the crowding out of printed newspaper readership is large, the effect of on-line is small, relative to some earlier predictions. A related study by Barbara Deleersnydera, Inge Geyskens, Katrijn Gielens, and Marnik G. Dekimpe (2002) used a data set of 85 Internet channel additions over 10 years in the

newspaper industries of the UK and The Netherlands. They concluded that the often-cited fears of cannibalization have been largely overstated in the information–goods industry. In this case the Internet need not disrupt established firms and channels.

Several related studies examine the consumer learning resulting in information cascading. J. Miguel Villas-Boas (2004) investigates how brand loyalty for the products generated by consumer learning affected her choices in the second period about competing products. Juanjuan Zhang (2009) demonstrates that observational learning and information sharing impact on consumer choices differently. Moreover, Welch (1992) examines the likelihood of cascades and optimal pricing in the market for initial public offerings. Banerjee (1992) devises a model for herd behavior as cascades. Bikhchandani, Hirshleifer, and Welch (1992) explain behavior conformity and the fragility of mass behavior by informational cascades. Hongbin Cai, Yuyu Chen, and Hanming Fang (2009) prove the inference effect via a field study, showing dining satisfaction increases by 13-20% when ranking information is presented to consumers. Several studies on consumer choice process in the Internet have been suggested. For example, Brynjolfsson and Smith (2001) find that consumers still prefer popular brands and are quite sensitive to ordinal ranking information even when the Internet eliminates the competitive advantage of products; so called "The Great Equalizer."

The primary contribution of this paper is to incorporate informational-cascading framework into a model to investigate the impact of Internet development on the printed newspaper industry. This is an especially interesting test in the contemporary environment, in which cascades have picked up speed as significant technological shifts offer consumers more comparative information.

The rest of the paper is organized as follows: Section 3 describes the research design; Section 4 presents the data. In Section 5, I introduce models to examine informational cascades according to various benchmarks. These models are estimated in Section 6, where I find that advancements in Internet technologies do not always threaten the survival of print media. Section 7 concludes the paper and suggests directions for future study.

3. Industry Background and Research Design

<u>Internet Usage in Korea</u>

To evaluate my hypothesis, I investigated Korea's Internet market since Korea has the ideal conditions available to study the actions of consumers who observed the choices of others.

Korea is one of the world's most wired countries: 65% of the population has access to the Internet and the nation has the fastest Internet connection in the world. In addition, the number of Internet searches per capita in Korea is higher than that of most other countries, which means there are also more Korean Internet users available to affect the choices of others (Graph 1).

Graph 1- Monthly Internet Searches Per Capita (By Country)
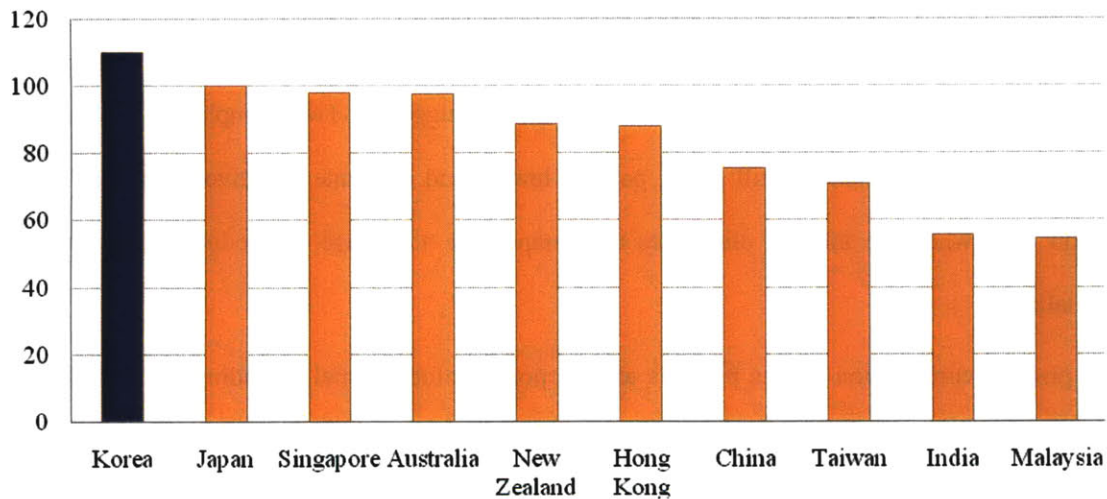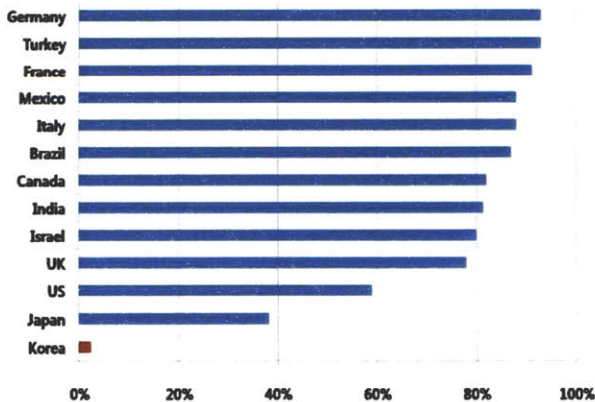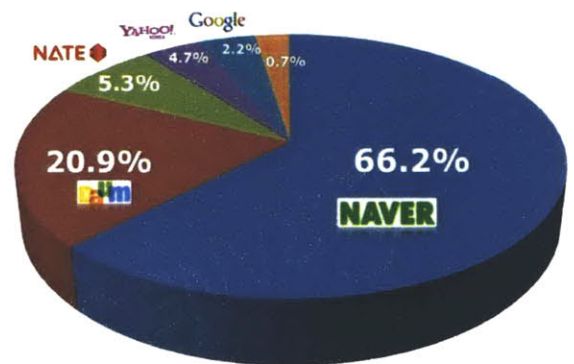

Source: ComScore, as of April 2009

NAVER.com

Although Google is the most frequently used Internet search site worldwide, national brands, particularly Naver.com, dominate the search engine market in Korea. Naver.com bears numerous similarities to Google. For example, if a user searches for news on Naver, he will be shown "Most Popular News" first, just as with Google News. Another similarity is that Naver categorizes news into World, Economy, Sports, Breaking (Spotlight) and Ranking (Top Stories). Further, Naver and Google both benefit from large numbers of users based on their strong search engine capabilities and the ability to provide ID-based services, such as Naver Mail and Gmail (Figure 1). As of October 2009, Naver holds 66.2% of the search engine market in Korea–three times greater market share than its closest competing search engine, Daum.net. Similarly, Google dominates market share worldwide: over 80% in most countries. However, in Korea, Google occupies only 2.2% of the search market space; one-thirtieth of Naver's presence. Similarly, Yahoo.com, one of the world's top search engines, has a negligible market share in Korea: 4.7% (Graph 2).

Graph 2- Google and Naver in Korea

A. Google's Worldwide Market Share by Country   *B. Market Share of Top Search Engine Companies in Korea*

Source: ComScore, as of October 2009

Figure 1 –Naver.com's "Most Popular News" and Google.com's "Top Stories" (*)

NAVER

"Most Popular News"

Google

"Top Stories"

(*) Domain: Naver News (http://news.naver.com) and Google News (http://news.google.com)

Business Newspapers

I examined business and economic-related newspapers because such periodicals are used by readers with similar demographics. There is a vital difference between other types of daily newspapers, such as those that provide political news and likely possess sharply opposite types of readerships. For example, *Chosun Ilbo* is a major newspaper in Korea. It has a daily circulation greater than two million. However, partly owing to its conservative editorial opinions, the paper's primary readership is composed of people over 40 years old. In contrast, readers 20-30 strongly prefer *Hankyure*, one of the most progressive newspapers.

Demographic similarity between readers of business newspapers is important because this study will examine the Internet's differential effect on two distinct groups: "top-journal" group (leaders) and "second-tier journal" group (followers). Since the two groups have similar monthly income and education, unobservable factors are expected to have the same effect on both groups, thereby making our coefficient of interest unbiased. In addition, I infer that business newspapers are less influenced by changes in the political environment, such as a regime change, which is a difficult variable to control. Therefore, I minimized the potential bias of the present quasi-experiment by solely analyzing the business newspaper group (Graph 3).

Graph 3 - Comparison of Demography between the "Leaders" and "Followers"
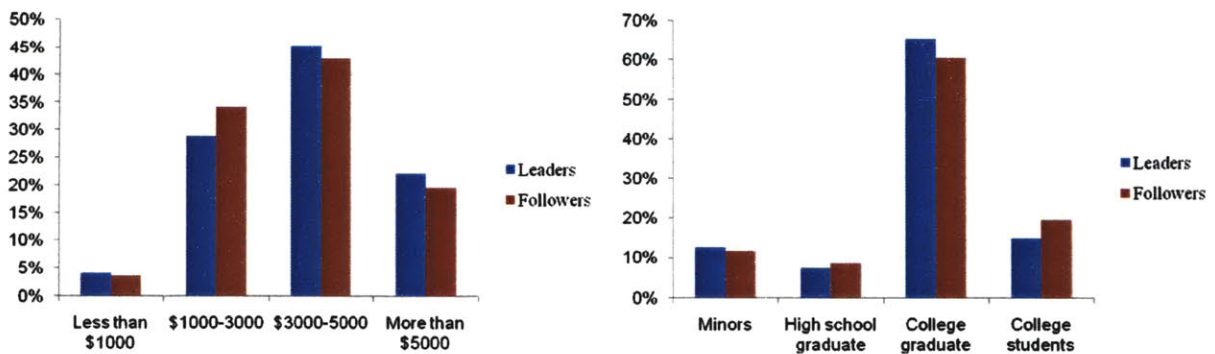
*Source: KoreanClick, 2006 Average

There are five business newspapers in Korea: *Maeil, Hankook, Herald, Financial News*, and the *Seoul Economic Newspaper*.[1] Among these I focused on *Maeil, Hankook, Herald*, and *Financial News*, since the *Seoul Economic Newspaper* does not offer data that is comparable to the other four newspaper companies.[2] Accordingly, I categorize *Maeil* and *Hankook* as "leaders" and the *Herald* and *Financial News* as "followers," based on their circulation.

Treatment: Policy Change at Naver.com

In my research, I empirically analyzed two critical events in Korea's Internet history and their effects on the development of on-line newspaper traffic. Naver.com, having changed two significant policies on the way in which it delivered news articles to consumers in 2006, presented a unique opportunity to test my hypothesis.

First, since December 2006, Naver.com has provided a personalized main-page editing service called "My News." Using this service, customers have the option to read only articles written by their preferred newspaper companies. In other words, users edit their own front page, selecting on-line newspapers they want to read from a list of 36 newspapers. Naver's "My News" service offers users more exposure to

---

[1] *Moneytoday*, the sixth ranked business newspaper in Korea, could be classified in the broad-based business newspaper category. However, in a strict sense, many Koreans do not regard *Moneytoday* as a conventional "business newspaper" given its on-line focus on its Web site as a primary source of revenue, versus circulation-based periodicals. Considering major printing and paper costs involved with business newspapers, accounting for nearly 40 % of revenues in 2008, I conclude that *Moneytoday* should be excluded from research subjects, which all demonstrate homogeneous characteristics, including primary sources of revenue and distribution channels.

[2] Unlike other economic newspapers, *Seoul Economic Newspaper* does not have its own Web site. Readers who want to view *Seoul Economic Newspaper* articles must via its parent company site, the *Korea Daily Newspaper* (www.hankooki.com). Therefore, it is not feasible to precisely separate *Seoul Economic Newspaper* readers from the entire number of visitors to the *Korea Daily Newspaper* homepage. In other words, there must be some visitors, who initially intend to read *Korea Daily Newspaper*, click on *Seoul Economic Newspaper* articles by chance, and vice versa. All Korean economic newspapers, except for *Seoul Economic Newspaper*, are purely specialized in economics and therefore do not cover general news. In addition, these newspapers run their own Web sites to feed economic articles directly to consumers.

news content provided by favorite newspapers. Therefore, "My News" service raises parameter α, the possibility that a consumer will choose her preferred on-line newspaper, and accelerates informational cascades discussed in Section 5. Moreover, on the Internet, we can easily observe the information goods choices of other users. Most Web portals provide rankings of the most popular news items ("Top Stories") on the company's homepage in almost real time.

Second, and even more important, search-engine technology has greatly improved to provide news research results that are relevant to the consumer. Before December 2006, if users clicked on the link to a story they were interested in reading among Naver.com search results, they were required to read the entire article inside Naver.com's domain page. In other words, Naver.com showed users the full text of another periodical's article from its own database (inlink technology). Therefore, reader traffic from all leading newspapers was counted as traffic to Naver.com itself, not to the on-line newspaper. This means that before December 2006, ranking in the "Most Popular News" section on Naver.com drastically underestimated the gap between leading newspapers and laggards.

However, after December 2006, if readers wanted to peruse an entire article, they could click through to a newspaper's homepage and end up there (outlink technology), not at Naver.com's site. Naver.com only shows such headlines and a few lines of a full story as indexed material and then transfers consumer traffic to the originating newspaper site, vastly increasing the number of clicks to leading newspaper Web sites. As a result, consumers are able to observe others' decisions and choices much more apparently via Naver's "Most Popular News" section. Through development of traffic trends on newspaper Web sites after 2007, I have observed the influence of the Internet on consumer preference and on newspapers. Consequently, thanks to the introduction of "outlink algorithm," the variance of noise σ^2 in the model (Section 5) will decrease much faster since consumers can identify other's preferences more clearly; therefore consumer choice will converge much faster towards equilibrium.

4. Data Description

The number of clicks data set, measured by the unique viewers (UV)[3] of on-line newspaper Web sites, is generated by KoreanClick (www.koreanclick.com). In Korea, several companies specialize in data compilation regarding Internet user behavior, including traffic changes on a particular Web site. Of these, I chose KoreanClick because it has been collecting stream data since July 2000, even before the Internet began to exert its influence on newspapers in earnest. Demographic information for this study is also gathered via KoreanClick. The company surveys a random sample of people who visit a certain Web site to determine user age, education, salary, job title and marital status.[4]

One of the primary findings of the present study results from examining the monthly time series data of the number of people who visit the two groups of newspaper Web sites. In Figure 2, I show a time path of the number of people who visit each newspaper's Web sites, measured by UV. From the figure, in the month of Naver's implementation of its "My News" service, the number of visitors increased noticeably. In November 2006, just before implementation of "My News," 631,011 people visited the *Herald* site, one of the "followers" periodicals; whereas 1,595,360 people visited *Maeil's* site, a periodical in the "leaders" group. Note that in December 2006, immediately following the implementation of "My News," the gap between the leading newspaper Web sites (*Maeil* and *Hankook*) and those of the followers (*Herald* and *Financial News*) began to widen substantially and continued increasing until May 2008– almost two years after the introduction of the "My News" service. This suggests that information cascades have picked up the speed as advances in Internet technologies offered consumers more and more accurate comparative information through the "Most Popular News" section on Naver.com.

---

[3] A Unique visitor (UV) is a statistic that describes a unit of traffic to a Web site, counting each visitor only once in the time frame of the report, typically one day.

[4] Other than a unique visitor, I also include page views (PV), new visits (NV) and the average frequency per day (AF) of all business newspaper company Web sites in Korea from January 2001 to present. Page views are the number of total visits to a site, which demonstrates the total number of clicks have occurred to an article from a specific Web site in the time frame of the report. In general, both UV and PV are important measurements on which to evaluate a Web site. Since UV measures the number of clicks, counting each visitor only once in the time frame of the report, I believe it is a more accurate measure of traffic. Therefore, I choose to use it for the present analysis.

Figure 2 –Monthly Number of Visitors to Business Newspaper Web sites (*)
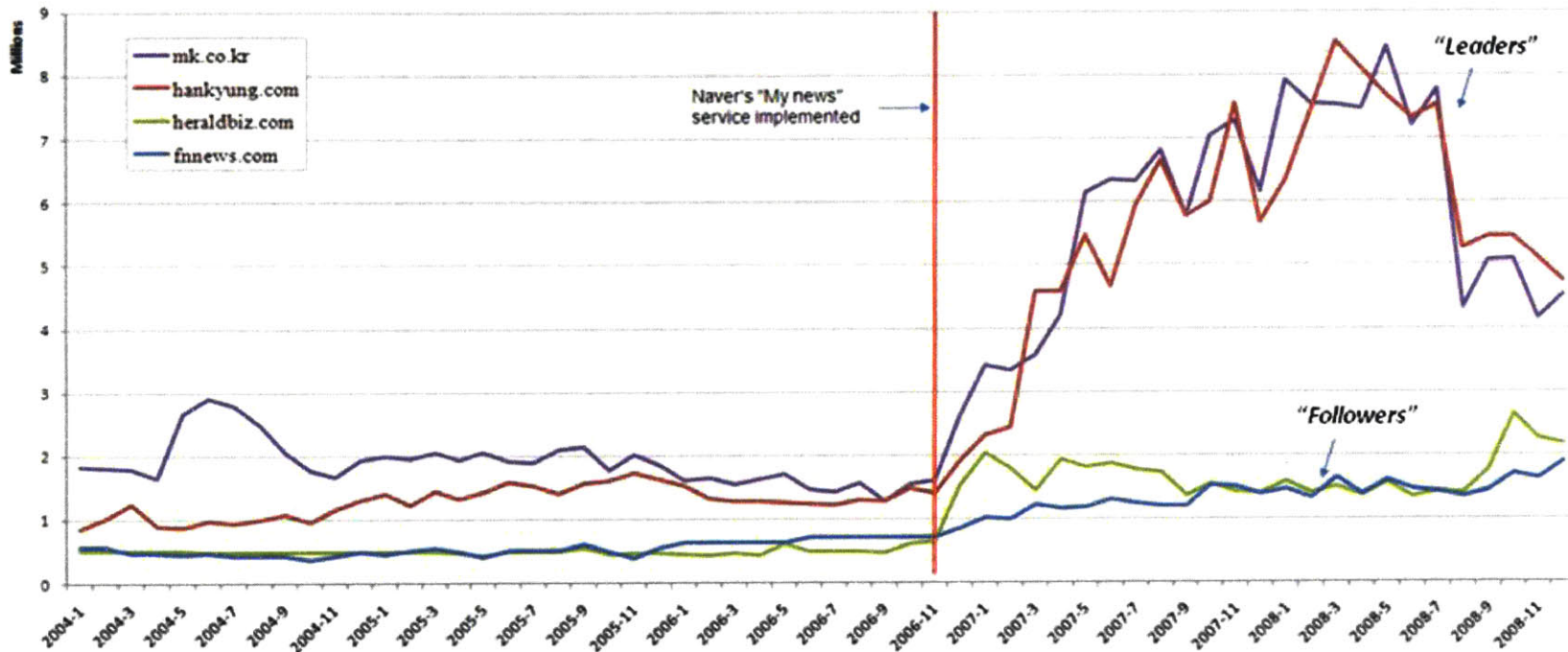
(*) Vertical axis represents the number of people who visit each site (unit: million UV), whereas the horizontal axis displays time.

The temporary decrease in traffic on mk.co.kr and hankyung.com that occurs since August 2008 is due to the fact that both newspapers temporarily stop providing news to Daum.net, which holds 20% share in the Korean search market. In June 2008, the movement to boycott American beef begins on some Web sites; The Internet forum Agora on Daum.net plays a central role in the on-line boycott. In particular, Internet activists launched a movement on Daum.net to boycott advertisers who place advertisements in those newspapers, including *Mail Business Newspaper* and *Hankook Economic Newspaper*, which had produced favorable articles on American beef imports. As a result, both newspapers stop feeding news to Daum.net, asking the search engine to take responsibility for providing space that allows the illegal activity of obstructing business. However, the newspapers resume news service on Daum.net in July 2009 and traffic on each newspaper Web sites rapidly recovers to pre-August 2008 levels before the sharp decline that follows, which is merely a temporary shock (For more details, see: http://www.sisapress.com/news/articleView.html?idxno=46683).

5. Model

In this section I use a model to examine informational cascades pattern in various benchmark settings where consumers choose information goods. For example, suppose there are two types of newspapers ("leaders" and "followers") and a population of size N. Proportion φ of the whole population prefers "leaders," while the remaining group prefers "followers," where $\frac{1}{2} < \varphi < 1$. Decision making in this model is sequential: one person chosen at random makes his decision first. With probability α, he will want to choose his preferred newspaper, but with probability (1-α), he will not be able to do so, in which case he will be exposed to either "leaders" or "followers," randomly, with a probability of $\frac{1}{2}$. In this case, the probability of the first player choosing "leaders" is as follows:

Pr (First player chooses the "leaders") = $P_1(\alpha) = \alpha \left(\varphi - \frac{1}{2} \right) + \frac{1}{2}$

Pr (First player chooses the "followers") = $1 - P_1(\alpha) = \alpha \left\{ (1 - \varphi) - \frac{1}{2} \right\} + \frac{1}{2}$

The next person, once again chosen at random,[5] makes her decision next and is also allowed to choose her preference with probability α. But, with probability (1-α), she will not be able to do so, in which case she must choose either "leaders" or "followers" by observing the previous player's choices. If I let $P_n(\alpha)$ be the probability of Nth player choosing the "leaders" group, then she will choose "leaders" with probability $P_1(\alpha)$, and choose "followers" with probability $1 - P_1(\alpha)$. The rest of the game proceeds similarly, with each randomly selected new player either allowed to choose her preference or make her decision on the basis of past player decisions.

---

[5] For simplicity, we assume that each person is chosen from the population *with replacement.*

1. No Internet

Before probing the effects of the Internet on print newspapers, let us first consider an extreme case wherein consumers cannot observe[6] others' choices, which is analogous to a world without the Internet. If there is no Internet and people cannot observe others' choices, then presumably, individuals will only choose their preferred news media; people who prefer "leaders" will read "leaders." whereas people who prefer "followers" will simply read "followers." Therefore, equilibrium is described simply as follows:

$$\text{Pr } (\text{N}^{\text{th}} \text{ player chooses "leaders")} = \varphi$$

$$\text{Pr } (\text{N}^{\text{th}} \text{ player chooses "followers")} = 1 - \varphi$$

2. Internet and Exogenous Preference

Now, suppose that one has Internet access and can observe the choices made by others. Using the Internet, consumers will no longer simply choose their preferences, because they are exposed to information from both groups, regardless of their preference. For example, even though an individual is extremely conservative and does not want to read liberal *New York Times* articles, he will certainly have no choice but to read these articles when searching for news on Google. Thus, the Internet exposes users to information to which they might not otherwise be exposed. In this case, the probability of Nth player choosing "leaders." $P_n(\alpha)$ can be derived as in Figure 3.

---

[6] It is evident that, without the Internet, consumers will be less likely to observe others' choices. Further, in many countries such as Korea, official sales quantities for each newspaper are not publicly available, which makes it impossible for consumers to observe or make purchase decisions based on other people's choices.

Figure 3 - N[th] Player's Choice Problem

Probability of each choice

Player observes the choices
made by the previous player

Player can choose her
preference

α

Player prefers
the "leaders"

φ

$\varphi\alpha$

$P_{n-1}(\alpha)$ Player clicks the $\varphi(1-\alpha)P_{n-1}$
"leaders"

Player cannot choose
her preference

$1-\alpha$ Player clicks the $\varphi(1-\alpha)(1-P_{n-1})$
"followers"

$1-P_{n-1}(\alpha)$

Nature randomly
chooses a person from
the population

$1-\varphi$

Player can choose her
preference

α

Player prefers
the "followers"

$(1-\varphi)\alpha$

$P_{n-1}(\alpha)$ Player clicks the $(1-\varphi)(1-\alpha)P_{n-1}$
"leaders"

Player cannot choose
her preference

$1-\alpha$ Player clicks the $(1-\varphi)(1-\alpha)(1-P_{n-1})$
"followers"

$1-P_{n-1}(\alpha)$

Thus, the probability of Nth player choosing "leaders." $P_n(\alpha)$ and "followers." $1-P_n(\alpha)$, respectively, are as follows:

Pr (n[th] player chooses the "leaders") = $P_n(\alpha) = \alpha(\varphi - P_{n-1}(\alpha)) + P_{n-1}(\alpha)$

Pr (n[th] player chooses the "followers") = $1 - P_n(\alpha) = \alpha\{(1-\varphi) - (1 - P_{n-1}(\alpha))\}$

Consequently, the probability of Nth player choosing "leaders" converges at φ[7]. Therefore, when personal preference is exogenous with a constant φ, the equilibrium choices of consumers converge to equilibrium choices without the Internet.

3. Internet and Endogenous Preference with Noisy Observation

Now suppose that players can still observe the choices made by previous players, but with a noise factor ε, where $\varepsilon \sim N(0, \sigma^2)$. Then, the probability that N[th] player will choose "leaders" will be $P_n(\alpha) + \varepsilon$, with the first player choosing "leaders" with probability of $\frac{1}{2} + \varepsilon$. Even in this case, as long as the noise has a

---

[7] With $0 < \alpha < 1$, $P_0(\alpha)=0.5$, φ will always greater than $P_{n-1}(\alpha)$, and $P_n(\alpha)$ will converge to α.

mean of 0 and has sufficiently small σ^2, information still draws individuals toward the same equilibrium action, although convergence will be slower (Bikhchandani et al., 1998).

4. Internet and Endogenous Preference

However, in the real world, individual preferences for information are affected by the media. In other words, exposed to information that they would not otherwise choose, individual preferences will change and, in this case, φ becomes endogenous. Although our model assumes that people are equally likely to be exposed by chance to either group ($P_0(\alpha) = \frac{1}{2}$), on average, consumers are more likely to read "leaders," since there is a fraction α of people who do not change what their preferences, even given exposure to the Internet. Since I chose individuals randomly from the population *with replacement* and people chosen are more likely to read "leaders" than "followers," (meaning $P_k(\alpha) > \frac{1}{2}$ when k $\geq$ 1), I observe increasing φ over time. Further, since I know that the probability of Nth player choosing "leaders" converges at φ, I can infer that the equilibrium choice of people measured by $P_n(\alpha)$ will approach 1 over time, which is the maximum possible value of φ.

$$P_n(\alpha) = \alpha\{\varphi(P_{n-1}(\alpha) - P_{n-1}(\alpha)\} + P_{n-1}(\alpha)$$

$$1 - P_n(\alpha) = \alpha\big[\{1 - \varphi(P_{n-1}(\alpha)\} - \big(1 - P_{n-1}(\alpha)\big)\big] + (1 - P_{n-1}(\alpha))$$

5. Link between the Model and Naver's "My News" Service

In December 2006 Naver.com introduced its "My News" service (i.e. personalized news listing service), affording consumers the option to choose their preferred on-line newspapers. In addition, the company changed its technological algorithms from an "inlink" to an "outlink" system, which made it possible for readers to observe the choices of others more clearly. The introduction of "outlink" brings the same effects as an increase in the total number the Internet users by uncovering the "hidden" (users can't notice

the actual number of clicks on a certain on-line newspaper) traffic under the past "inlink" system. The change is a general impact from the Internet. In contrast, "My News" is not adopted by every search engine and content aggregator like Google and Yahoo. However, consumers are becoming increasingly loyal to information they prefer by receiving daily news content from their preferred newspaper through e-mail. In my model, the effect of "My News" service will increase α (the probability of choosing a preferred on-line newspaper), whereas the "outlink" system will decrease the variance of noise σ^2. Since two changes in Naver.com's policy result in a higher α and a smaller σ^2, consumer choice will converge faster toward equilibrium than before, which provides excellent insight into the way the Internet reshapes the geography of our news outlets.

6. Estimation

In this section, I examine the responses of a number of visitors to the "My News" service in Naver.com. The predicted service effect is to increase α and to decrease σ^2 in the previous model, thereby reinforcing the information cascades and disproportionately increasing the number of visitors to leaders and followers newspapers. To evaluate the responses of a number of visitors, I use the regression discontinuity design (RDD) with the key assumption that the number of visitors during the month just before the "My News" service starts is a valid counterfactual for the number of visitors during the first month of the service.

I estimate a specification of the discontinuity in trend of log number of visitors of the following form, separately for leaders and followers groups.

$$\ln UV_{it} = \beta_0 + \delta_1 post_t + \delta_2 post_t f(t) + f(t) + \Pi X_{it} + \rho_i + \varepsilon_{it} \qquad (1)$$

ln UV_{it} is the log of the number of people who visit the Web site of newspaper i in month t; X_{it} is a vector of covariates that includes sales revenue of each newspaper company, GDP growth, and other political factors, while ρ_i represents a newspaper fixed effect. Post is a variable indicating the period after implementation of the "My News" service in December 2006. The function f(t) reflects overall time trends in the number of visitors, which I linearly parameterize with a slope varying in the pre- and post-service periods, using Post × f(t), the interaction between Post and linear time trend, which captures the changing slope of the linear time trend after implementation of the service. Particular coefficients of interest are δ_1 and δ_2. δ_1 captures the average difference in the number of visitors, while δ_2 captures the average difference in linear time trends between the pre- and post-service periods. I am interested in δ_2 as well as in δ_1 since I predict that the information cascades will be reinforced as consumers repeatedly interact on the Internet and catch each other's preferences.

Table 1 – Log UV Response to Naver's "My News" Service

	Log UV of "Followers"			Log UV of "Leaders"		
	(1)	(2)	(3)	(1)	(2)	(3)
Post	1.093***	0.803***	0.758***	1.309***	0.565**	0.565***
	(0.0505)	(0.221)	(0.219)	(0.053)	(0.289)	(0.260)
Post × f(t)		-0.000	0.002		0.014***	0.0142***
		(0.005)	(0.005)		(0.006)	(0.006)
Linear Time Trend		0.010***	0.008***		0.002	0.002
		(0.003)	(0.003)		(0.003)	(0.002)
Newspaper Fixed Effect			Yes			Yes
Observations	81	81	81	120	120	120
R^2	0.8559	0.8790	0.8839	0.8338	0.8449	0.8753

Note: Robust standard errors are in parentheses. Post is a variable indicating the period after implementation of the "My News" service in December 2006. The variable Post × f(t) is the interaction between Post and linear time trend, which captures the changing slope of the linear time trend after implementation of the service.

Table 2 – Response of Log Share of the "leaders" to Naver's "My News" Service

	Log Share of "Leaders"		
	(1)	(2)	(3)
Post	0.0437***	-0.0857**	-0.0857**
	(0.00854)	(0.0433)	(0.0433)
Post × f(t)		0.0046***	0.0047***
		(0.0010)	(0.0010)
Linear Time Trend		-0.0030***	-0.0031***
		(0.0004)	(0.0004)
Newspaper Fixed Effect			Yes
Observations	240	240	240
R^2	0.0989	0.2420	0.2420

Note: Robust standard errors are in parentheses.

*** Significance at 99% level, ** Significance at 95% level, * Significance at 90% level

Table 1 displays results from estimating the regression equation (1). In Column (2), I show estimated discontinuity in the number of visitors, controlling only for linear time trend. In the specification shown in Column (3) of Table 1, I include newspaper fixed effect and other time-varying covariates to find that estimated discontinuity is stable when adding these controls. Specifically, note column (2) of each group. The UV for "leaders" continues to increase much more rapidly than that of "followers" after the "My News" service is introduced. The rate of change (Post×f(t)) for leaders is much steeper than for followers, thus I find a statistically significant informational cascades effect and this phenomenon is reinforced through the inference effect. Figure 4 shows the fit of the linear time trend to average actual number of visitors. The linear time trend seems successful at fitting the time profile.

Table 2 shows that proportions of UV shares leaders and followers change after the introduction of the "My News" service.   In Column (1), the UV share of leaders group increases by approximately 4.4%. In addition, I find that UV share of leaders increase by an additional 1% every month owing to Naver.com's "My News" service by controlling linear time trend and using fixed effect. These results are shown in Column (2) and (3) respectively, and the results are statistically significant.

Figure 4 – Predicted Versus Actual Average Number of Visitors (*)

     A, "Followers" Group: Herald Business (heraldbiz.com) and Financial News (fnnews.com)

     B, "Leaders" Group: Maeil Business (mk.co.kr) and Hankook Economic (hankyung.com)
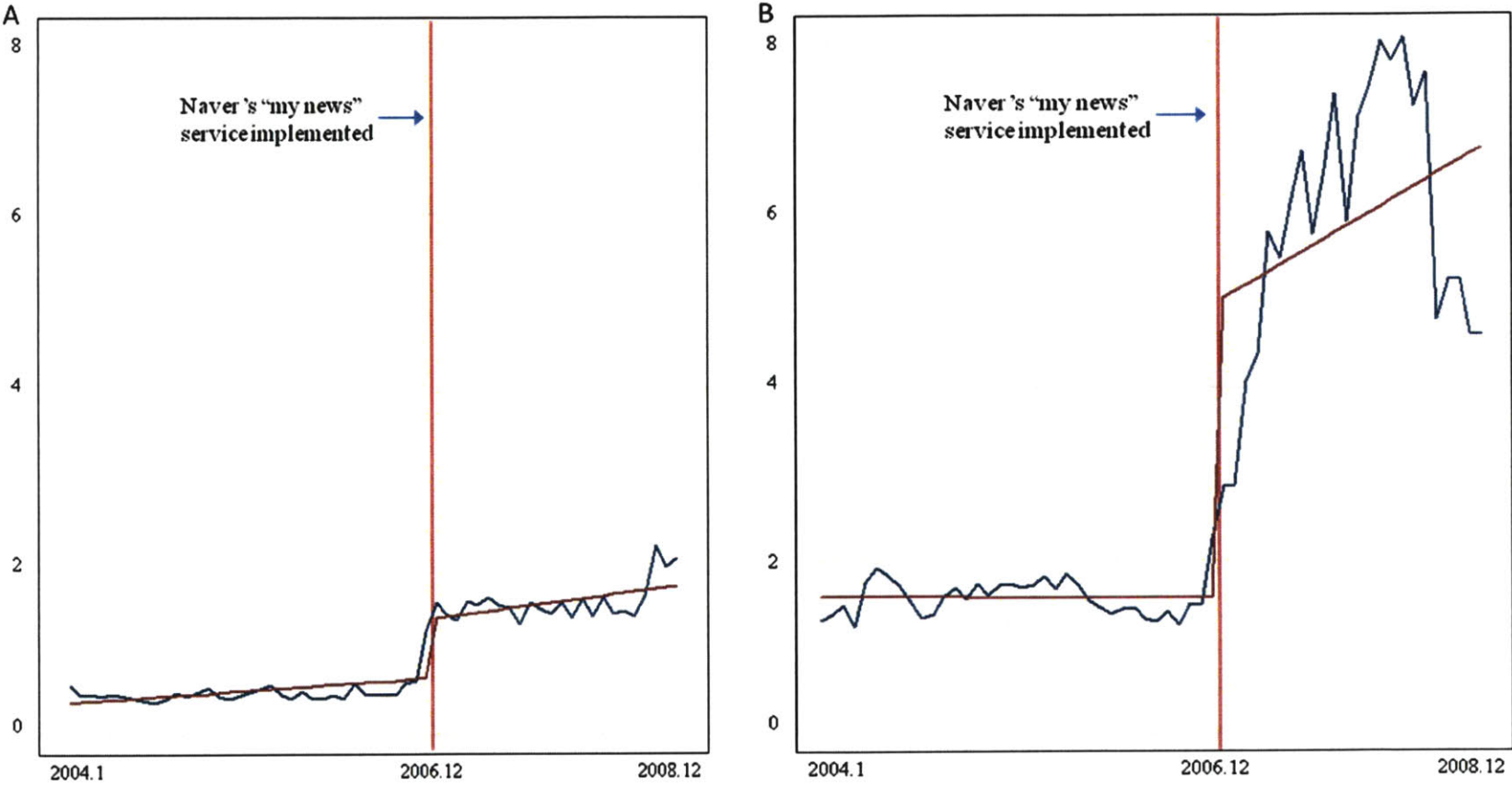
(*) Vertical axis represents the number of people visited (unit: million UV), while horizontal axis includes time trends.

7. Conclusion and Discussion

In this paper I presented the influence of Internet development on consumer behavior and choice in the Korean newspaper market, representative of information goods. My research offers empirical analysis and examples that show consumer behavior in regard to information consumption: Consumers observe and take notice of the behavior and choice of others when choosing their favorite media company.

By analyzing a unique click data set for the past nine years, I found that traffic on Web sites of leading newspapers had increased significantly compared to those of followers as Internet technology development offered consumers more accurate comparative information through the "Most Popular News" section. Therefore, users can now easily and clearly observe the choices of others, as well as their own preferences, which may generate inference effects and future informational cascades. The increased growth in traffic to leading Web sites subsequently boosts the exposure of their news to the "Most Popular News." This event creates a virtuous cycle where increased exposure helps to raise brand awareness of leading newspapers and subsequently attracts more visitors. In this way, advances in Internet technology offers unique opportunities to leading media organizations that have greater resources to charm more consumers. Therefore, my research provides evidence that technology development in the Internet is not a "foe" that paints a grim future for printed newspapers. Rather, it can benefit market leaders the most and provide them with new opportunities to establish brand image and loyalty.

These findings have important implications for the U.S. newspaper industry. In other words, such consumer behavior analysis is worth attention because it suggests new opportunities for traditional newspapers which are suffering through what could be their worst financial crisis since the Great Depression.

Moreover, it has long been a practice that printed newspapers do not make their circulation volumes public. Readers have no way of obtaining accurate information about market share–or popularity–of each newspaper. Accordingly, readers choose information goods without knowing what fellow consumers have

chosen. However, with the advent of the Internet, real-time news ranking provides readers with a popularity standard, reinforcing the informational cascades.

The study also provides empirical evidence of a theoretical model of informational cascading (Banerjee 1992, Bikhchandani, Hirshleifer, and Welch 1992). Lastly, it can be inferred from this research that the common claim YouTube is a problematic replacement for television stations, might not be true. From these perspectives, my research suggests several interesting areas for future research. For example, can advances in Internet technology help leading newspapers generate more advertisement revenue, improving revenue structure? Can leading newspapers enjoy cross-selling effects by using an image enhanced by the Internet? Can this popularity indirectly help leading newspapers generate additional revenue, if not directly via advertisements due to the recent intensification of competition in the market? The last two questions could be answered by comparing the profit change associated with the educational service offered by the leading newspapers and the followers, just as the *Washington Post* saw a sales growth of 51% last year through Kaplan, one of its subsidiaries in the education sector.

References

Banerjee, Abhijit V., "A simple model of herd behavior," *The Quarterly Journal of Economics*, Vol.107, No.3, 1992, p797-817

Bikhchandani, Sushil, David Hirshleifer, and Ivo Welch, "A theory of fads, fashion, custom, and cultural change as informational cascades," *Journal of Political Economy*, Vol.100, No.5, p992-1026

Bikhchandani, Sushil, David Hirshleifer, and Ivo Welch, "Learning from the behavior of the others: conformity, fads, and informational cascades," *The Journal of Economic Perspectives*, Vol.12, No.3, 1998, p151-170

Brynjolfsson, Erik and Michael D. Smith, "The Great Equalizer? Consumer Choice Behavior at Internet Shopbots," 2001, MIT Sloan Working Paper No. 4208-01

Cai, Hongbin, Yuyu Chen, and Hanming Fang, "Observational Learning: Evidence from a Randomized Natural Field Experiment," *American Economic Review*, 2009. Forthcoming

Deleersnydera, Barbara, Inge Geyskens, Katrijn Gielens, and Marnik G. Dekimpe, "How Cannibalistic is the Internet Channel? A study of the Newspaper Industry in the United Kingdom and The Netherlands," *International Journal of Research in Marketing*, Vol. 19, Issue 4, 2002, p 337-348

Filistrucchi, Lapo, "The Impact of Internet on the Market for Daily Newspapers in Italy," European University Institute working paper series, No. ECO2005/12

Gentzkow, Matthew, "Valuing New Goods in a Model with Complementarity: On-line Newspapers," *The American Economic Review*, Vol.97, No.3, 2007, p713-744

Hahn, Jinyong, Petra Todd, Wilbert Van der Klaauw, "Identification and estimation of treatment effects with a regression-discontinuity design," *Econometrica*, Vol.69, No.1, p201-209

Kaiser, Ulrich, "Is a Newspaper's Companion Web site a Competing Outlet Channel for the Print Version?," 2005, ZEW - Centre for European Economic Research Discussion Paper No. 05-078

Kirchhoff, Suzanne M., "The U.S. Newspaper Industry in Transition," Congress Research Service Report for Congress, July 8, 2009

Oberholzer-Gee, Felix and Anita Elberse, "Superstars and Underdogs: An Examination of the Long Tail Phenomenon in Video Sales," 2007 HBS Working Paper Series, No. 07-015

Patterson, Thomas E., "The Internet and the Threat It Poses to Local Media: Lessons from News in the Schools," 2007. The Carnegie-Knight Task Force on the Future of Journalism Education

Salganik, Matthew J., Peter Sheridan Dodds, and Duncan J. Watts, "Experimental Study of Inequality and Unpredictability in an Artificial Cultural Market," *Science*, 2006, 311, 854-856

Tucker, Catherine and Juanjuan Zhang, "How Dees Popularity Information Affect Choices? A Field Experiment," 2009, MIT Sloan Research Paper No.4655-07

Villas-Boas, J. Miguel, "Consumer Learning, Brand Loyalty, and Competition," *Marketing Science*, Vol. 23, No. 1, 2004, 23, p134-145

Zhang, Juanjuan, "The Sound of Silence: Observational Learning in the U.S. Kidney Market," *Marketing Science*, 2009. Forthcoming