HARVARD | BUSINESS | SCHOOL

Do Friends Influence Purchases in a Social Network?

Raghuram Iyengar
Sangman Han
Sunil Gupta

Working Paper

09-123

Do Friends Influence Purchases in a Social Network?

Raghuram Iyengar
Sangman Han
Sunil Gupta[1]

February 26, 2009

[1] Raghuram Iyengar (riyengar@wharton.upenn.edu) is Assistant Professor at the Wharton School, University of Pennsylvania, Philadelphia, PA 19104; Sangman Han (smhan@skku.edu) is Professor of Marketing at the Sung Kyun Kwan University, Korea; and Sunil Gupta (sgupta@hbs.edu) is the Edward W. Carter Professor of Business Administration at the Harvard Business School, Soldiers Field, Boston, MA 02163.

Do Friends Influence Purchases in a Social Network?

Abstract

Social networks, such as Facebook and Myspace have witnessed a rapid growth in their membership. Some of these businesses have tried an advertising-based model with very limited success. However, these businesses have not fully explored the power of their members to influence each other's behavior. This potential viral or social effect can have significant impact on the success of these companies as well as provide a unique new marketing opportunity for traditional companies.

However, this potential is predicated on the assumption that friends influence user's behavior. In this study we empirically examine this issue. Specifically we address three questions – do friends influence purchases of users in an online social network; which users are more influenced by this social pressure; and can we quantify this social influence in terms of increase in sales and revenue.

To address these questions we use data from Cyworld, an online social networking site in Korea. Cyworld users create mini-homepages to interact with their friends. These mini-homepages, which become a way of self-expression for members, are decorated with items (e.g., wallpaper, music), many of which are sold by Cyworld. Using 10 weeks of purchase and non-purchase data from 208 users, we build an individual level model of choice (buy-no buy) and quantity (how much money to spend). We estimate this model using Bayesian approach and MCMC method.

Our results show that there are three distinct groups of users with very different behavior. The low-status group (48% of users) are not well connected, show limited interaction with other members and are unaffected by social pressure. The middle-status group (40% users) is moderately connected, show reasonable non-purchase activity on the site and have a strong and positive effect due to friends' purchases. In other words, this group exhibits "keeping up with the Joneses" behavior. On average, their revenue increases by 5% due to this social influence. The high-status group (12% users) is well connected and very active on the site, and shows a significant negative effect due to friends' purchases. In other words, this group differentiates itself from others by lowering their purchase and strongly pursuing non-purchase related activities. This social influence leads to almost 14% drop in the revenue of this group. We discuss the theoretical and managerial implications of our results.

Do Friends Influence Purchases in a Social Network?

Social networks have become a cultural phenomenon. Facebook, one of the largest social networking sites in the U.S. was founded in 2004. By February 2009, it boasts more than 175 million active users and continues to grow rapidly. Worldwide these users spend 3.0 billion minutes each day on Facebook. More than 850 million photos and 5 million videos are uploaded on the site each month.[2] There are hundreds of other similar sites including Myspace, Friendster, Xanga and Bebo. This cultural and technological revolution is not limited to the United States. Myspace has already launched its international sites in Britain, Australia and France and plans to expand its services to nine other countries in Europe and Asia in the near future. More than 70% of Facebook users are outside the U.S. and more than 35 translations are available on the site. Other countries have their own versions of Facebook and Myspace. For example, Cyworld, which started before Myspace and Facebook were conceived in the US, had over 21 million registered users in South Korea by mid-2007, or approximately 40% of the South Korean population. It has over 90% penetration in the 20-29 year old market. Cyworld users upload about 50,000 videos and 5 million photos every day.

In spite of this cultural and social revolution, the business viability of these social networking sites remains in question. While many sites are attempting to follow Google and generate revenues from advertising, there is significant skepticism if advertising will be effective on social networking sites. Seth Goldstein, co-founder of SocialMedia Networks, recently wrote on his Facebook blog that a banner ad "is universally disregarded as irrelevant if it's not ignored entirely," (New York Times, Dec 14, 2008). Recognizing this, in November 2007, Facebook experimented with a new program called

[2] Source: http://www.facebook.com/press/info.php?statistics , accessed February 23, 2009

Beacon, which shared purchases of a friend with a user with the hope that this would be viewed as “trusted referral” and generate more sales for its advertisers. The program backfired due to privacy issues but Facebook asserted that it would continue to evaluate this kind of program. Cyworld has been selling music and other virtual items (e.g., wallpaper) to its users for many years with the belief that friends influence each other’s purchases of these items.

If friends indeed influence purchases of a user in a social network, it could potentially be a significant source of revenue for the social networking sites and their corporate sponsors. The purpose of this study is to empirically assess if this is indeed true. Specifically, we wish to answer the following questions:

- Do friends influence purchases (frequency and/or amount) of a user in a social network?
- Which users are more influenced by this social pressure?
- Can we quantify this social influence in terms of percentage increase in sales revenue?

We address these questions using a unique data set from the Korean social networking site, Cyworld. Using the actual (rather than reported or surveyed) data of over 200 users for several months, we build a model to examine how friends influence the purchases of a user. We estimate this model using Bayesian methods which provide us parameter estimates at an individual user level.

Our results show that there is a significant and positive impact of friends’ purchases on the purchase probability of a user. Even more interestingly, we find that there are significant differences across users. Specifically, we find that this social effect is zero for 48% of the users, negative for 12% of the users and positive for 40% of the users.

Further examination reveals systematic differences across these user groups. Users who have limited connection to other members are not influenced by friends' purchases. However, positive social effect is observed in moderately connected users. These users exhibit "keeping up with the Joneses" behavior. On average, this social influence translates into a 5% increase in revenues. In contrast to this group, highly connected users show a negative effect of contagion. To maintain distinctiveness, these users tend to reduce their purchases of items when they see their friends buying them. This negative social effect reduces the revenue for this group by more than 14%. We discuss the reasons and implications of these findings.

The paper is organized as follows. We begin with a brief description of related literature to put our research in context. Next, we describe the data since a clear understanding of the data is helpful in developing the model. The model and its estimation are discussed next, followed by results and conclusion.

RELATED LITERATURE

Research on social networks has captured the effect of social influence on consumers' purchase decisions across a variety of contexts. Such an effect has been variously termed as bandwagon effect (Leibenstein 1950), peer influence (Duncan, Haller and Portes 1968; Manski 1993, 2000), neighborhood effect (Bell and Song 2007; Case 1991; Singer and Spilerman 1983), conformity (Bernheim 1994), and contagion (Van den Bulte and Lilien 2001; Iyengar, Van den Bulte and Valente 2008). Recent work has also considered how social influence can operate even within a retail context (Argo, Dahl and Morales 2006, 2008).

Across these studies, typically two approaches have been used for characterizing the network among consumers – spatial proximity and self-report. For example, Bell and Song (2007) capture the effects on potential customers of an online grocery retailer due to exposure to spatially proximate existing customers. This has much precedence in both the marketing and sociology literature (Case 1991; Singer and Spilerman 1983). Iyengar, Van den Bulte and Valente (2008) use the social network among physicians elicited through self reports and show that there is a positive contagion effect at work in physicians' decisions to adopt a new drug. The use of self reports also has much precedence in the sociology literature (Coleman, Katz and Menzel 1966; Valente et al. 2003). Both these methods, however, have limitations. The geography based method, while being objective, involves the contagion to be inferred i.e., other alternative explanations such as spatial heterogeneity, spatial autocorrelation have to be carefully tested. The self report measure is direct but suffers from all the typical survey related biases such as selective memory and social desirability.

Data from online social networks directly give detailed information about how consumers interact with the rest of the network without any of the above mentioned weaknesses. For instance, on Cyworld, members set up mini-home pages that they use to display pictures, play their favorite music, record their thoughts and decorate with their chosen virtual items (e.g., wallpaper). The site provides information about users purchase as well non-purchase activities.

Much past work using online social networks has explored the role of network structure on the diffusion of information in the social network. Some of this work has emphasized the existence of power laws in degree distribution (Barabási 2002; Barabási and Albert 1999; Barabási and Bonabeau 2003) and have called attention to highly connected nodes in networks or hubs. See Newman (2003) for a review of the role of network structure for many processes such as product adoption occurring over the network. Keller and Barry (2003) showed that people who influence others tend to have relatively large numbers of social links, and Gladwell (2000) described these people as "connectors." These connectors have mega-influence on their neighbors, because they are linked with a large number of people. Weimann (1994) provides an overview of the research on opinion leaders across many contexts of this nature.

Some recent work has questioned the influence of such hubs. Watts and Dodds (2007), based on simulation studies, report that large cascades of information diffusion are not driven by hubs but by a critical mass of easily influenced individuals. In contrast, Goldenberg, Han, Lehmann and Hong (2009) provide evidence that the success or failure of information diffusion does depend upon the adoption decision of social hubs. They, however, differentiate innovator hubs from follower hubs and show that while innovator hubs are important in initiating the diffusion, it is the follower hubs that are important in determining the size of diffusion.

Recent research has used online social network data to address several questions. Trusov, Bucklin and Pauwels (2008) compare the effect of customer invitations to join the network (word-of-mouth marketing) with traditional advertising. Using a time-series methodology, they show that word-of-mouth marketing has a substantially larger carry over effect than traditional marketing. Trusov, Bodapati and Bucklin (2009) examine a member's activity (specifically the count of daily logins) on a social network as a function of both self-effects and the activity level of his/her friends. Using a Poisson model for daily logins, they identify specific users who most influence others' activities. We complement these studies by examining the impact of social influence on *actual purchase behavior* and quantify these effects in revenue terms.

As this brief review indicates, few past studies have focused on purchase behavior within a social network. The focus of our study is on empirically testing whether purchases in the social network are contagious.

Do Friends Help or Hinder Purchase?

Past research has documented that consumers have a need to differentiate themselves from others (Ariely and Levav 2000; Snyder and Fromkin 1980; Tian, Bearden and Hunter 2001). Consumers' tastes, which include their purchasing behavior, attitude and preferences they hold, can signal their social identity (Belk 1988; Douglas and Isherwood 1978; Levy 1959; Wernerfelt 1990) and can be used by others to make desired inferences about them (Calder and Burnkrant 1977; Holman 1981; McCracken 1988; Muniz and O'Guinn 2001). While tastes do signal social identity, what others infer from one's choice depends upon group membership (Berger and Heath 2007; McCracken 1988; Muniz and O'Guinn 2001). For example, Berger and Heath (2007) find that people may converge or diverge in their tastes based on how much their choice in a given context signals their social identity. They discuss the example of the adoption of Harley motorcycles and note that if many tough people ride Harley motorcycles, then Harleys may signal a rugged identity. However, if suburban accountants start adopting Harleys as well, then the meaning of adopting a Harley might become diffuse. This is the standard fashion cycle (Bourdieu 1984; Hebdige 1987; Simmel 1971), and is a problem faced by many major luxury brands such as Louis Vuitton and Burberry (Han, Nunes and Drèze 2008). Within a social network, we can potentially observe such convergence and divergence of tastes. It is also possible that these effects vary across users.

DATA

Our data comes from Cyworld, a Korean social networking company. Cyworld was started in 1999 by a group of MBA students from the Korean Institute of Science and Technology. Initially called People Square, it was quickly renamed Cyworld. “Cy” in Korean means relationship, which defined the goal of the company. By mid-2007, Cyworld had 21 million registered users in a country of about 50 million people.

Users create their mini-home pages (called minihompy in Cyworld), which they use to display pictures or play their favorite music. These mini-home pages also contain bulletin boards on which users can record their thoughts and feelings. Users take great pleasure in decorating their own home pages by purchasing virtual items such as furniture, household items, wallpaper, as well as music. A mini home page is seen by users as a means for self expression, and virtual items enable users to achieve this goal. In 2007, Cyworld generated $65 million or almost 70% of its revenue from selling these items. The remaining revenue was generated from advertising and mobile services.

In addition to purchasing virtual items, members also engage in non-purchase related activities. Members regularly update the content (pictures, diaries, music, etc) of their own mini-home pages and visit the homepages of their friends to keep abreast of their updated content. If a user finds some content on a friend’s mini-home page interesting, she can “scrape" it from friend’s page onto her own mini-home page. The scraping function has the effect of replicating what members find interesting, thus generating a viral effect and increasing the value of network for all members. Each mini-home page is able to record and display the number of visitors it receives, the replies to messages posted there, and the content scraped onto other mini-homepages. These feedback measures serve as indices of popularity.

One attractive feature of Cyworld is that it offers members the opportunity to designate certain other members as “first neighbors” - a designation not unlike best friends. Members can list their existing friends as first neighbors and make new friends with whom they establish first neighbor ties. Finally, Cyworld also gives members the ability to search outside of their first-neighbor networks by means of a function called first-neighbor waves, which allows individuals to search the networks of their first neighbors.

The dataset for this study is a log file of 640 panelists, who agreed to install pc-meters on their computers to allow tracking of their online navigation behavior. The log file contains such information as duration and page views of the main categories such as main page, mini-home page, club, gift shop, and submenus of each category. The log file also includes relational data such as frequency of replies, uploading, and number of pages seen. These data were collected from September 20 to December 8, 2004. From these 640 panelists, we selected 208 members who are fully connected, i.e. no one is isolated from the rest of the members. We use data from these 208 connected members for this study.

MODEL AND VARIABLES

Each week, a user decides whether or not to buy virtual items from Cyworld. Although Cyworld sells a large number of these items, the data are fairly sparse for any particular item. Therefore we combine all items into a single category and focus on buy-no buy decision for any of these items. If a user decides to buy, she needs to decide how much money to spend on these items. These two decisions of the user (choice and quantity) are modeled as follows (Krishmamurthi and Raj 1988).

Choice Model

The decision to buy an item depends on user-specific characteristics (e.g., her past behavior) as well as influence of other members. Within a social network, a member's status and his influence can be defined and measured in several different ways. Rogers and Cartano (1962) discussed three ways: (1) self-designation, i.e., asking survey respondents to report to what extent they perceive themselves to be influential, (2) sociometric techniques, i.e., computing network centrality scores after asking survey respondents whom they turn to for advice or information or after observing interactions through other means (e.g., citations among scientists), and (3) the key informant technique where selected people are asked to report their opinion about who are the key influential members. Whereas self-designation is the most popular technique among marketing academics (e.g., Childers 1986; Myers and Robertson 1972), the sociometric technique has been more popular among social network analysts (e.g., Coleman, Katz and Menzel 1966; Valente et al. 2003). This approach is especially suited for an online social network, where we have easy access to information on members' interactions.

Specifically, we define the utility for user *i* from making a purchase in week *t* as:

$$U_{it} = \alpha_{0i} + \alpha_{1i}(Indegree)_{it-1} + \alpha_{2i}(Outdegree)_{it-1} + \alpha_{3i}(Social\ Influence)_{it-1} + \alpha_{4i}(Past\ Purchase)_{it-1} + \alpha_{5i}(Indegree * Social\ Influence)_{it-1} + \alpha_{6i}(Outdegree * Social\ Influence)_{it-1} + \varepsilon_{it}^{c}$$

where the covariates are defined as follows.

Indegree is the number of members, within our sample, who visit a particular member in a given week. Indegree is the most basic measure of status or prestige of a member in a network (Wasserman and Faust 1994). Popular or prestigious members have a large following of people who constantly visit their homepages to learn about the latest trends or news. Users who have low indegree may be inclined to buy items to gain popularity among friends, while users who are already popular (high indegree) may want to buy items to retain their status. It is also possible that popular users may avoid buying commercially available items to ensure that they remain unique. Indegree varies across members and time (within a member on a weekly basis). We use a member's indegree from last week as a covariate that may influence her purchase in the current week.[3]

Outdegree is the number of members, within our sample, visited by a specific member in a particular week. Outdegree reflects the proclivity of a member to scan and interact with her network of friends. A member with high outdegree has more opportunities to be influenced by her friends' purchases. At the same time, the need for uniqueness may drive this person to avoid buying items that her friends have already

[3] Lagged terms avoid endogeneity problems, unless (1) people are forward-looking not only about their own behavior but also that of others *and* (2) social ties over which influence flows are symmetric. The first condition is quite unlikely in large networks, and the second condition does not hold in our data. Lagged terms also allow us to break free of the reflection problem (Manski 1993, 2000).

purchased. Similar to the indegree measure, this covariate also varies across members and within a member on a weekly basis, and we use the lagged term in the utility equation.

Social Influence. While indegree and outdegree provide indirect measures of social influence, a more direct measure is the actual purchase behavior of friends. We operationalize this direct social influence as the exposure of a particular member *i* to other members' purchases through his weekly visit behavior using lagged endogenous autocorrelation terms (Strang 1991). The extent to which member *i* is exposed in week *t* to prior purchases is captured through the term $\Sigma_j\, w_{ijt-1}\, z_{j,t-1}$ where w_{ijt-1} is 1 if member *i* visits member *j* in week *t-1,* 0 otherwise; and $z_{j,t-1}$ is the amount of money spent by member *j* in last week.

Note that $\Sigma_j\, w_{ijt-1}$ is the outdegree of member *i*. If we define $z_{j,t-1}$ as a 0-1 variable based on whether or not member *j* bought an item in the last week, then social influence variable is strictly less than or equal to the outdegree of a member. In other words, this construct would suggest that a member gets influenced largely from friends who made a purchase last week. To allow for the possibility that a friend who has bought several items may have more influence than a friend who has bought fewer items, we define $z_{j,t-1}$ as the money spent by member *j* in last week.

To assess the extent to which sociometric measures (indegree and outdegree) moderate the effect of this social influence variable, we also use two interaction terms.

Past Purchase. We use a member's own lagged weekly monetary value of purchases to capture a member's proclivity to buy.

Table 1 provides summary statistics for these covariates for our dataset.

Insert Table 1 here

Quantity Model

Quantity or the amount of money spent on items by user *i* in week *t*, conditional on buying in a given week, is defined in a similar way. Specifically,

$$\begin{aligned}\log\ Q_{it} &= \beta_{0i} + \beta_{1i}(Indegree)_{it-1} + \beta_{2i}(Outdegree)_{it-1} \\ &\quad + \beta_{3i}(Social\ Influence)_{it-1} + \beta_{4i}(Past\ Purchase)_{it-1} \\ &\quad + \beta_{5i}(Indegree * Social\ Influence)_{it-1} + \beta_{6i}(Outdegree \\ &\quad * Social\ Influence)_{it-1} + \varepsilon_{it}^{Q}\end{aligned}$$

We assume that $\boldsymbol{\varepsilon}_{it} = \{\varepsilon_{it}^{C}, \varepsilon_{it}^{Q}\}$, has a bivariate normal distribution i.e., $\boldsymbol{\varepsilon}_{it} \sim N(0, \Sigma)$. For identification purposes, the variance of the utility function (Σ_{11}) is set to 1. We estimate both the covariance between the choice and the quantity random shocks (Σ_{12}), and the variance of the quantity random shock (Σ_{22}).

So far, we have developed the model for a member *i*. Next, we specify the heterogeneity across members. Let $\boldsymbol{\alpha}_i = \{\alpha_{0i}, \alpha_{1i}, ..., \alpha_{6i}\}$, $\boldsymbol{\beta}_i = \{\beta_{0i}, \beta_{1i}, ..., \beta_{6i}\}$ and let $\boldsymbol{\delta}_i = \{\boldsymbol{\alpha}_i', \boldsymbol{\beta}_i'\}$. We assume that $\boldsymbol{\delta}_i$ is normally distributed in the population, i.e., $\boldsymbol{\delta}_i \sim \mathrm{N}(\mu, \Lambda)$. This completes our model specification.

We also estimate two null models. Null Model 1 does not contain the main and the interaction effects of the social influence variable. Null Model 2 includes only the main effects of social influence. A comparison of these two null models with our full model will help in better understanding the effect of social influence variable. For both null models, we specify customer heterogeneity similar to that in our full model.

We estimate our model and the two null models using Bayesian approach and MCMC methods. For each model, we obtain parameter draws based on 100,000 iterations after a burn-in period of 50,000 iterations of the MCMC chain.

RESULTS

Model Comparison

For model comparison, we use these draws to calculate the log-marginal likelihoods (LML) for each model.   Low absolute values denote a better model. The LML for Null Model 1 (No Contagion) is -1268.94, for Null Model 2 (Main Effect) is -1263.45 and for the full model is -1253.52. These numbers can be used to calculate the log Bayes Factors (LBF). In comparison to Null Model 1, the LBF for Null Model 2 is 5.49 (=1268.45-1263.94) and that for the full model is 15.42 (=1268.94-1253.52). According to Kass and Raftery (1995), if the LBF among two models is greater than 5, then it shows strong support for the model with the lower absolute LML. Thus, the full model is best supported by the data. A comparison of the log-marginal likelihoods shows that including social influence in the model is important and so is the inclusion of the interaction terms between sociometric measures of indegree/ outdegree and the social influence. Next, we present the parameter estimates for the full model.

Parameter Estimates

The parameter estimates for the full model are given in Table 2. The table presents estimates for the *population means* of the parameters. The numbers in parenthesis are the 95% posterior intervals around the mean and the significant posterior means are shown in bold.

Insert Table 2 Here

Four main things emerge from these results. First, indegree, outdegree and contagion are significant in the choice model. In the quantity model, all parameters other than the intercept are insignificant. Second, the main effect of indegree and outdegree is significant and negative, suggesting that popular members have a need for uniqueness which drives them to buy less. This finding is consistent with past work within online communities (Han and Kim 2008). Third, the social influence variable has a strong and positive impact on members' choice decision. In other words, on average, friends' purchases in the past have a strong positive impact on a member's current purchase. Fourth, we find that the covariance between choice and quantity errors (Σ_{12}) is not significantly different from zero, indicating no latent correlation among these decisions, perhaps due to a lack of explanatory power of the quantity model. Further, we find that the variance of the quantity error (Σ_{22}) is 0.48.

In addition, Bayesian estimation approach allows us to estimate parameters at an individual level, so we can see the heterogeneity of these parameters across our sample of users. Figures 1-3 show the histograms of *individual-level parameters* for indegree, outdegree and social influence in the choice model. These figures show a significant heterogeneity in how individual members get influenced by social factors.

Who Gets Positively or Negatively Influenced by Friends and Why?

While it is possible to interpret parameters (indegree, outdegree, social influence etc.) for each individual, it is much easier and insightful to see how these parameters interact to influence the overall purchase probability of members. This approach not only provides us a net effect of all the variables but also offers us the magnitude of this effect.

We achieve this by running the following simulation based on the estimated individual-level parameters. We simulate the data for all 208 members for 10 weeks. Further, to incorporate the uncertainty in the individual-specific parameters, we generate the data for the simulation concurrently with our estimation algorithm. For each member and a sample of her individual-specific parameters from the MCMC chain, we generate a total of 200 paths, where a path for a member represents her weekly buy/no buy decision and the associated monetary value for 10 weeks. The algorithm for generating one path is as follows.

For the first week, we initialize all variables to their actual values from the data. We then generate the utility associated with the buy/no buy decision and simulate the monetary value of the purchase if a member decides to purchase. This monetary value is used to create the lagged monetary variable for the subsequent week. To generate the social influence variable, we use the actual weekly visit patterns among members coupled with their simulated lagged monetary value. This process is iterated over 10 weeks. We generate 200 such paths for each member and each sample of her individual-specific parameters. We then average over all 200 paths to obtain the probability of purchase at every week. Finally, we incorporate the uncertainty in the individual-specific parameters by averaging the probabilities over all sampled values of the parameters from the MCMC chain.

For computing the impact of the social influence variable, we simulate two datasets according to the data generation process described above. The first dataset is simulated using our full model while the second dataset is generated without the social influence variable. We then compare the difference in the probability of purchase for all 208 members across the 10 weeks.

We categorize members in three groups – those with positive change in purchase probability due to social influence, those with negative change in purchase probability, and those with no effect. We find that 84 members or 40% of our sample has a positive impact due to social influence (see Table 3). The average increase in the purchase probability of this group is 0.01, which translates into a 5.3% increase in revenue. About 12% of our sample or 25 members reduce their probability of purchase due to their friends' purchase activity in the previous period. On average, this represents a decline of purchase probability by 0.05, which translates into a more than 14% drop in the revenue for this group. For the remaining 99 members (about 48% of sample), there is no significant impact on purchase. These marked differences across members are masked in an aggregate analysis.

Understanding the Three Groups

Who are the people in each of these groups? Table 3 shows the average indegree and outdegree of these three groups. These results suggest that the group that is influenced negatively by friends' purchases had high indegree and outdegree compared to the other groups. In other words, these 12% of the people represent the “high status” people. The group with “no effect” has the smallest indegree and outdegree, i.e. it is the least connected and engaged group. To further corroborate this group-level analysis, we ran a simple regression with change in purchase probability for an individual member as the dependent variable and her indegree and indegree-square as the two independent variables. We find a marginally significant positive coefficient for indegree ($p=0.056$) and a significant negative coefficient for indegree-square ($p < 0.001$). This confirms that there is an inverted-U relationship between indegree and change in purchase probability. Similar result was found for outdegree.

Insert Table 3 Here

To further understand the behavior of these three groups, we examine their non-purchase related activities on the social network. Table 4 presents a list of activities for which data are available to us for all the members in our sample. Table 5 presents the means of these activities for the three user groups. Means which are significantly different from the other groups are indicated in bold.

Insert Tables 4 and 5 Here

Table 5 shows that the members in the segments differ in a systematic manner in both their purchase and non-purchase activities. In particular, we find that members with zero social effect show little activity; members with positive effect have an intermediate level of activity while members with negative effect have very high level of activity. We also note that, on several activity measures, the low status group is significantly different (lower) than the middle status group, which in turn is significantly different (lower) than the high status group.

These findings reveal an interesting characterization of members when viewed from the perspective of maintenance of status and need for differentiation (Bourdieu 1984; Berger and Heath 2007; Van den Bulte and Joshi 2007). The group with zero effect contains members who are not well connected to other members as well as show little non-purchase related activity. These are essentially members who play no active role in the online community and are of the lowest status as they visit few members and very few visit them.

The group with positive effect constitutes members who are positively affected by other members making purchases in their neighborhood. They try to maintain their status by making their own purchases as they fear that not doing so might undo their status (Burt 1987). This is the typical “keeping up with the Joneses” effect such that the utility of these consumers is affected not just by what they do but also by what others do in their neighborhood (Abel 1990; Campbell and Cochrane 1999).

Finally, the group with negative effect contains well connected, high status, members – they visit several members and many visit them. These members show high level of non-purchase activity and their probability of purchase is lowered if other members around them are purchasing. These two findings indicate that, as other members around them imitate their purchases to gain status, these high status members further differentiate themselves by pursuing non-purchase related activity. For example, they seem to maintain their status by uploading their own unique content. This finding is similar to work on characteristics of opinion leaders or the elite in the fashion industry, who tend to abandon one type of fashion and adopt the next in order to differentiate themselves from the masses (Simmel 1971).

Our finding is also consistent with the middle-status conformity thesis in sociology (e.g., Philips and Zuckerman 2001), which suggests that member segmentation should be in three tiers - low, middle and high status. Across these three segments, it proposes that low-status people do not imitate others because they feel that it will not help them gain more status. High status people do not imitate others very much because they feel quite confident in their own judgment and the legitimacy of their actions. It is only middle-status people who feel that social pressure for the fear of falling in the social ranks.

Our study not only empirically confirms these theories but also provides the size of these groups (48% low status with zero effect, 40% middle status with positive effect, and 12% high status with negative social effect). We further assess the revenue impact of social influence. Specifically, we find that in our sample social influence reduces revenue for the high status members by about 14%, while it increases revenue for middle-status members by about 5%.

Managerial Implications

Increasing clutter in traditional advertising medium (e.g., TV), higher usage of recording devices such as TiVo, fragmentation of consumers, and increasing use of the Internet especially among younger consumers, has led marketers to start experimenting with alternative forms of communication. One of the promising, yet less well understood, forms is viral marketing. For instance, Proctor and Gamble operates Tremor and Vocalpoint, two word-of-mouth marketing services, to promote many of its products. Social networking sites, such as Facebook and Myspace, have reported significant growth in their membership but at the same time are struggling to find a sustainable business model. The advertising-based model, that worked so well for Google, has had limited success at social networking sites since users come to these sites to interact with their friends and not to search or buy products.

Our study points to a promising area for the social networking sites as well as for the large advertisers, such as P&G or Sony. If the purpose of advertising is to make consumers aware of the product and create interest among potential users, then it is possible for Sony to achieve the same result by giving its, say, new digital cameras free to the high status consumers. Similarly, music companies can offer free songs to this group of users. In many cases, offering free products to the right group of people may in fact be cheaper than traditional advertising. Our study shows that presence of these items among consumers can have a strong and positive social effect among middle-status members. Using the methodology offered in our study, the sponsoring company can also identify the size of different groups and the likely impact on appropriate metrics such as brand awareness or sales.

CONCLUSIONS

In this paper, we started with three questions: (a) do friends influences purchases of a user in a social network? (b) which users are more influenced by this social pressure? And (c) what is the impact of this social influence in terms of changes in sales and revenues. To address these questions, we developed a choice and quantity model that captures the effect of social influence on a member's decision to purchase. We used customer-level weekly data from CyWorld and Bayesian methodology to estimate the model.

We found significant heterogeneity among users. Our results show three distinct user groups: a) Low status members (48% in our sample), who are not well connected to other members, experience little or no social effect and hence do not change their purchase patterns due to friends' purchase behavior, b) Middle status members (40% in our sample), who are moderately conected, and show a strong positive effect when their friends buy items and c) High status members (12% in our sample), who are the most well connected, but show a negative social effect.

To understand how members strive for differentiation, we linked the purchase related activity of members with their non-purchase related activity. The group with negligible contagion effect contained members who are not well connected to other members as well as show little non-purchase related activity. The group with positive contagion effect constitutes members who exhibit a moderate level of non-purchase activity. They try to maintain their status by primarily making purchases as they fear that not doing so might undo their status. This is the typical "keeping up with the Joneses" effect. Finally, the group with negative effect contains well connected, high status members. These members show a very high level of non-purchase activity and their probability of purchase is lowered if other members around them are purchasing. This is consistent with the typical fashion cycle wherein opinion leaders or the elite in the fashion industry tend to abandon one type of fashion and adopt the next in order to differentiate themselves from the masses. As other members around them imitate their purchases to gain status, these high status members further differentiate themselves by pursuing non-purchase related activity.

We also quantify the social influence in terms of changes in purchase probability and revenues. Our results show that middle-status users show, on average, a 5% increase in revenue due to social influence. In contrast, the high status group's revenue declines by almost 14% due to these social effects.

Our findings are relevant for social networking sites and large advertisers. The members in high status group have an influence on those in the middle status group for the diffusion of a new product. However, a successful diffusion in the middle status segment may make high status members lose interest in the new product. This interplay of product diffusion and customer segmentation leaves much room for future research.

Table 1: Summary statistics of the data

Variable	Mean	Standard Deviation	Minimum	Maximum
Weekly Purchase Incidence	0.20	0.40	0.00	1.00
Weekly Monetary Value	0.03	0.09	0.00	1.32
Social Influence	0.03	0.10	0.00	1.25
Indegree	0.85	1.15	0.00	9.00
Outdegree	0.85	1.61	0.00	13.00

Note: Number of users = 208; Number of observations =2080

We scale the total monetary value by 10,000 Korean Wons.

Table 2: Parameter estimates for the proposed model

Parameter	**Choice Model**	**Quantity Model**
Intercept	**-1.03**	**-2.35**
	(-1.23, -0.84)	(-2.67, -2.03)
Indegree	**-0.22**	0.06
	(-0.44, -0.02)	(-0.18, 0.29)
Outdegree	**-0.26**	0.07
	(-0.51, -0.02)	(-0.18, 0.33)
Social Influence	**2.78**	-0.75
	(0.19, 5.52)	(-3.24, 1.95)
Past Purchase	-0.22	-0.69
	(-1.63, 0.99)	(-1.90, 0.53)
Indegree*Social Influence	-0.65	0.06
	(-2.49, 1.19)	(-1.77, 1.97)
Outdegree*Social Influence	-1.18	0.33
	(-2.67, 0.09)	(-1.04, 1.65)

Note: The table presents the estimates for the population means of the parameters. The numbers in parenthesis are the 95% posterior intervals around the mean and the significant posterior means are shown in bold. For the model, we scale the total monetary value by 10,000 Korean Wons.

Table 3: Characterization of members in the three segments

Variables	Positive Social Effect	Zero Social Effect	Negative Social Effect
Difference in probability of purchase due to social influence	0.01	0.00	-0.05
% Change in revenue due to social influence	5.3	0	-14.1
Indegree	0.76	0.66	**1.88**[b]
Outdegree	**1.10**	**0.21**[a]	**2.51**[b]
Number of members	84	99	25
% of sample	40	48	12

Notes: Positive Effect contains members for whom the probability difference > 0.001.

Zero Effect contains members for whom the probability difference < 0.001 and > -0.001.

Negative Effect contains members for whom the probability difference < -0.001.

a: This indicates a significant different ($p < 0.05$) between the low status and middle status groups.

b: This indicates a significant difference ($p < 0.05$) between the high status and middle status groups.

Significant differences across groups are highlighted in bold.

Table 4: Description of non-purchase related activity measures

Variable	Description
Scrap	The number of scraping activities from others' mini-homepage
Reply	The number of replying to others' mini-homepage
Page	Total number of page views at other's mini-homepage
Duration	Total duration time at other's mini-homepage
Replied	The number of replies received by others
Upload	The number of uploading activities at his/her own mini-homepage

Table 5: Non-purchase activities of members in the three segments

Variables	Positive Social effect **(Mid-Status)**	Zero Social effect **(Low-Status)**	Negative Social effect **(High-Status)**
Scrap	1.67	0.62	**4.20** [b]
Reply	**3.29**	**0.48** [a]	**10.60** [b]
Page	**78.50**	**8.45** [a]	**213.20** [b]
Duration	3526.37	**312.86** [a]	6844.48
Replied	2.10	1.69	**9.80** [b]
Upload	32.43	**13.08** [a]	50.48
Number of members	84	99	25

Notes:

a: This indicates a significant different ($p < 0.05$) between the low status and middle status groups.

b: This indicates a significant difference ($p < 0.05$) between the high status and middle status groups.

Significant differences across groups are highlighted in bold.

Figure 1: Histogram of Individual-specific Indegree Parameter

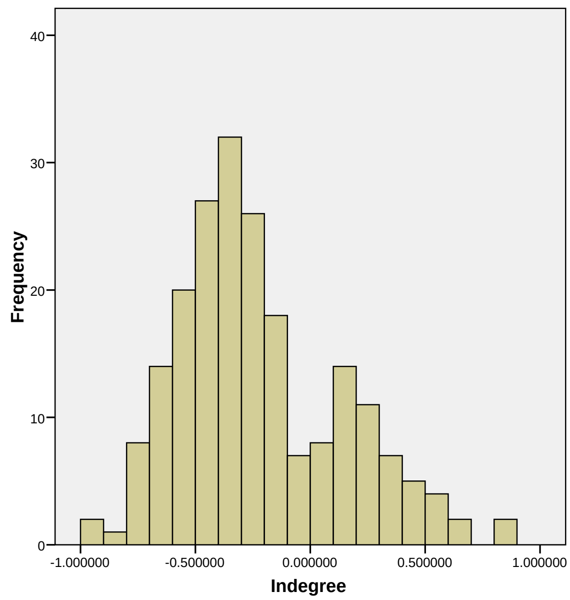


Figure 2: Histogram of Individual-specific Outdegree Parameter

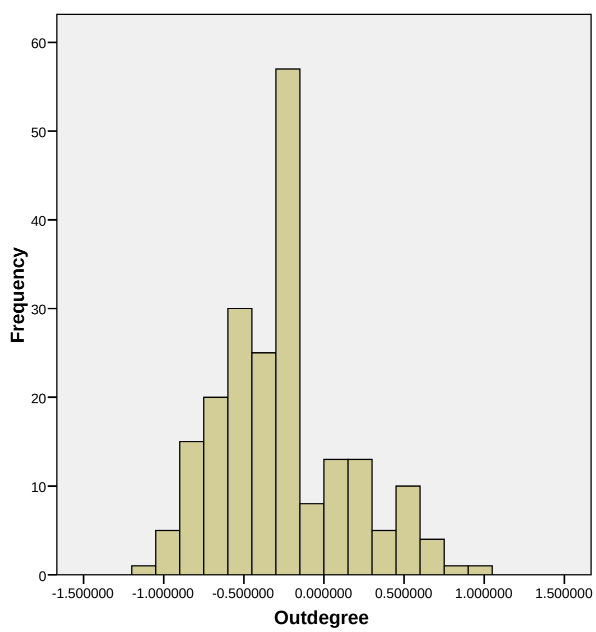
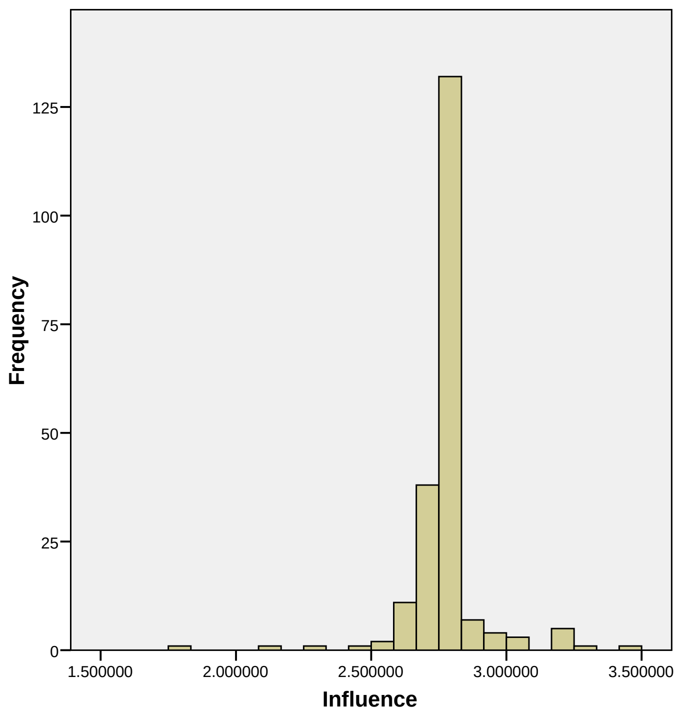


Figure 3: Histogram of Individual-specific Social Influence Parameter



REFERENCES

Abel, Andrew B. (1990), “Asset prices under habit formation and catching up with the joneses*,” American Economic Review Papers and Proceedings,* 80, 38-42.

Argo, Jennifer J., Darren W. Dahl and Andrea C. Morales (2006), “Consumer contamination: how consumers react to products touched by others,” *Journal of Marketing*, 70 (2), 81-94.

Argo, Jennifer J., Darren W. Dahl and Andrea C. Morales (2008), “Positive consumer contagion: responses to attractive others in a retail context,” *Journal of Marketing Research*, 45 (6), 690-701.

Ariely, Dan and Jonathan Levav (2000) “Sequential choice in group settings: taking the road less traveled and less enjoyed,” *Journal of Consumer Research*, 27 (3), 279-290.

Barabási, Albert-Lázló (2002), *Linked: The New Science of Networks,* Perseus Publishing.

Barabási, Albert-Lázló and Reka Albert (1999), “Emergence of scaling in random networks,” *Science*, 286, 509-12.

Barabási, Albert-Lázló and Eric Bonabeau (2003), “Scale-Free Networks,” *Scientific American,* 288(5), 50-59.

Belk, Russell W. (1988), “Possessions and the extended self,” *Journal of Consumer Research*, 15 (2), 139-168.

Bell, David R. and Sanyoung Song (2007), “Neighborhood effects and trial on the internet: evidence from online grocery retailing,” *Quantitative Marketing and Economics,* 5 (4), 361-400.

Berger, Jonah and Chip Heath (2007), “Where consumers diverge from others: identity signaling and product domains,” *Journal of Consumer Research*, 34 (2), 121-134.

Bernheim, Douglas B. (1994), “A theory of conformity,” *Journal of Political Economy,* 102 (5), 841-877.

Bourdieu, Pierre (1984), “*Distinction: A Social Critique of the Judgment of Taste*,” Harvard University Press, Cambridge, MA.

Burt, Ronald S. (1987), “Social contagion and innovation: cohesion versus structural equivalence,” *American Journal of Sociology*, 92, 1287–1335.

Calder, Bobby J. and Robert E. Burnkrant (1977), “Interpersonal influence on consumer behavior: an attribution theory approach,” *Journal of Consumer Research*, 4 (1), 29-38.

Campbell, John Y. and John H. Cochrane (1999), "By force of habit: a consumption-based explanation of aggregate stock market behavior," *Journal of Political Economy*, 107 (2), 205-251.

Case, Anne C. (1991), "Spatial patterns in household demand," *Econometrica*, 59 (4), 953-965.

Childers, Terry L. (1986), "Assessment of the psychometric properties of an opinion leadership scale," *Journal of Marketing Research*, 23 (2), 184-188.

Coleman, James S., Elihu Katz and Herbert Menzel (1966), *Medical Innovation: A Diffusion Study*, New York: Bobbs-Merrill.

Douglas, Mary and Baron Isherwood (1978), *The World of Goods: Towards an Anthropology of Consumption*, New York: Norton.

Duncan, O. Dudley, Archibald O. Haller, and Alejandro Portes (1968), "Peer influences on aspirations: reinterpretation," *American Journal of Sociology*, 74 (2), 119–137.

Gladwell, Malcolm (2000), *The Tipping Point,* New York: Little, Brown and Company.

Goldenberg, Jacob, Sangman Han, Donald R. Lehmann and Jaewon Hong (2009), "The role of hubs in the adoption process," *Journal of Marketing*, 73(March), 1-13.

Han, Sangman and BeomJun Kim (2008), "Network analysis of an online community," *Physica A: Statistical Mechanics and its Applications,* 387 (23), 5946-5951.

Han, Young Jee, Joseph C. Nunes and Xavier Drèze (2008), "First impressions: status signaling through brand prominence," Working paper, University of Pennsylvania.

Hebdige, Dick (1987), *Subculture: The Meaning of Style*, London: Routledge.

Holman, Rebecca H. (1981), "The imagination of the future: a hidden concept in the study of consumer decision making," *Advances in Consumer Research*, 8 (1), 187-191.

Iyengar, Raghuram, Christophe Van den Bulte and Thomas W. Valente (2008), "Opinion leadership and social contagion in new product diffusion," Working paper, University of Pennsylvania.

Kass, Robert E., and Adrian Raftery (1995), "Bayes factors," *Journal of the American Statistical Association*, 90, 773-795.

Keller, Ed and Jon Barry (2003), *The Influentials: One American in Ten Tells the other Nine How to Vote, Where to Eat, and What to Buy,* New York: The Free Press.

Krishnamurthi, Lakshman and S.P. Raj (1988), “A model of brand choice and purchase quantities price sensitivities,” *Marketing Science*, 7, 1, 1-20.

Leibenstein, Harvey (1950), “Bandwagon, snob, and veblen effects in the theory of consumers’ demand,” *Quarterly Journal of Economics*, 64 (2), 183-207.

Levy, Sidney J. (1959), “Symbols for sale,” *Harvard Business Review*, 33 (March–April), 117–124.

Manski, Charles F. (1993), “Identification of endogenous social effects: the reflection problem,” *Review of Economic Studies*, 60 (3), 531-542.

Manski, Charles F. (2000), “Economic analysis of social interactions,” *Journal of Economic Perspectives*, 14 (3), 115-136.

McCracken, Grant (1988), *Culture and Consumption: New Approaches to the Symbolic Character of Consumer Goods and Activities*, Bloomington: Indiana University Press.

Muniz, Albert M. Jr. and Thomas C. O’Guinn (2001), “Brand community,” *Journal of Consumer Research*, 27 (March), 412-432.

Myers, James H. and Thomas S. Robertson (1972) “Dimensions of opinion leadership,” *Journal of Marketing Research*, 9 (1), 41-46.

Newman, Mark E. J. (2003), “The structure and function of complex networks,” *SIAM Review*, 45, 167-256.

Philips, Damon J. and Ezra W. Zuckerman (2001), “Middle-status conformity: theoretical restatement and empirical demonstration in two markets,” *American Journal of Sociology*, 107 (September), 379-429.

Rogers, Everett M. and David G. Cartano (1962), “Methods of measuring opinion leadership,” *Public Opinion Quarterly*, 26 (Autumn), 435-441.

Simmel, Georg (1971), “Fashion.” Donald N. Levine, ed. *Georg Simmel on Individuality and Social Forms*. University of Chicago Press, Chicago, IL, 294–323.

Snyder, C. R. and Howard L. Fromkin (1980), *Uniqueness: The pursuit of human difference*, New York: Plenum Press.

Singer, Burton H. and Seymour Spilerman (1983), “The representation of social processes by Markov models,” *American Journal of Sociology*, 82 (1), 1–54.

Strang, David (1991), “Adding social structure to diffusion models: an event history framework,” *Sociological Methods and Research*, 19 (February), 324-353.

Tian, Kelly Tepper, William O. Bearden and Gray L. Hunter (2001), “Consumers' need for uniqueness: scale development and validation,” *Journal of Consumer Research*, 28 (1), 50-66.

Trusov, Michael, Randolph E. Bucklin and Koen Pauwels (2008), “Effects of word-of-mouth versus traditional marketing: findings from an internet social networking site,” *Journal of Marketing,* Forthcoming.

Trusov, Michael, Anand Bodapati and Randolph E. Bucklin (2009), “Determining influential users in internet social networks,” Working paper, University of Maryland.

Valente, Thomas W., Beth R. Hoffman, Annamara Ritt-Olson, Kara Lichtman, and C. Anderson Johnson (2003), “Effects of a social-network method for group assignment strategies on peer-led tobacco prevention programs in schools,” *American Journal of Public Health*, 93 (November), 1837-1843.

Van den Bulte, Christophe and Yogesh V. Joshi (2007), “New product diffusion with influentials and imitators,” *Marketing Science,* 26 (3), 400-421.

Van den Bulte, Christophe and Gray L. Lilien (2001), “Medical innovation revisited: social contagion versus marketing effort,” *American Journal of Sociology*, 106 (5), 1409-1435.

Wasserman, Stanley and Katherine Faust (1994), *Social network analysis, methods and application*, Cambridge Press.

Watts, J. Duncan and Peter S. Dodds (2007), “Influentials, networks, and public opinion formation,” *Journal of Consumer Research, 34 (4), 441.*

Weimann, Gabriel (1994), *The Influentials: People Who Influence People*, Albany, NY: State University of New York Press.

Wernerfelt, Birger (1990), “Advertising content when brand choice is a signal,” *Journal of Business*, 63 (1), 91-98.