Who are the customer evangelists, and what are they worth? A network model to measure customer referral value

Abstract: In this paper we propose a model to assess the value of customer referral activity. Customer referral value (CRV) represents the economic value a customer is likely to generate through recommendations to others. Previous research shows that this metric is a good proxy of customers' likelihood to attract new customers through recommendation. This article proposes a new methodology to estimate CRV. We first model social influence through a metric called personal network exposure, which is calculated from a social network model. Then we check whether this metric is a significant predictor of adoption using event history analysis in the statistical validation. In a third step, we estimate how each customer contribute to the personal network exposure of those customer who adopted after him, and finally we assess CRV as the sum of the product between the likelihood that the influential activity of a focal customer affects each customer that adopts after him, multiplied by these clients' life time value.

A novel data set, which serves to test the model, combines the dynamics of social relationships among a set of customers over ten years with their consumption patterns. Compared with other methodologies that measure customers' recommendation value, we show that our approach is more robust, flexible, and easy to implement, and it accurately reflects the role of social interactions in influencing the likelihood from customers' recommendations.

Guillermo Armelini (ESE – Universidad de los Andes - Chile)
Erica Salvaj (FEN-UDD –Chile)
Julian Villanueva (IESE – Universidad de Navarra – Spain)

1. INTRODUCTION

The practice of identifying and targeting "evangelists", customers committed with the firm that are willing to talk positively about its products to a potential prospect, is becoming very popular among companies. Notice for example the successful case of Vocalpoint, a word of mouth marketing program of Procter & Gamble designed to encourage moms, clients of the consumer goods company, to talk positively about P&G brands. More than 600,000 women that have been carefully recruited taking into consideration the size of their social network, actively participate in the program, allowing P&G to reach a huge amount of consumers at a very effective cost. Other companies, from different industries, such as Cingular, United Airlines and Sprint PCS (Kumar et al. 2007; Ryu and Feick 2007) are also motivating current customers to bring in new clients through different kinds of recommendation programs. This trend in the marketing strategy of many firms is not surprising, as researches have showed that referrals are one of the most important sources of gaining new customers (Brown and Reingen 1987; Helm 2003; Murray 1991; Reingen and Kernan 1986), making "evangelists" an attractive marketing medium.

However, while there is a consensus among scholars about the positive influence of customers' recommendation in consumer behavior, (Godes et al. 2005; Villanueva et al. 2008), little is known about how firms can indentify customers who are willing to spread the word out and how much value they create for the firm because of their advocacy.

One way to identify these customers is analyzing the referral activity of the so-called influentials or opinion leaders, "special" individuals that exert a disproportionate amount of influence on the behavior of "others" (Summers 1970; Watts 2007). The main idea is that influentials can serve as important seed points in diffusion processes (Corey 1971; Engel et al. 1969), accelerating the rate of adoption by other persons through their recommendations and social influence (Valente and Davis 1999). Yet recent studies question the effectiveness of this strategy, either because it is difficult to identify opinion leaders (Iyengar et al. 2009), because the process of WOM, in most cases, is only minimally driven by influentials (Watts and Dodds 2007), or even worst, because the real influence of influentials has never been defined accurately. As Duncan Watts stated (2007:206):

"the claim that influentials (however defined) are important (in some sense) because they (somehow) spread their options (defined in some manner) throughout some (unspecified) network is a statement that sounds like it says something, but in fact relies on so many unstated and ambiguous assumptions that it says nothing in particular at all"

Alternatively, instead of finding and targeting opinion leaders, it might be interesting to identify the main "talkers" estimating customers' referral activity with a financial metric. Regarding this, marketing scholars suggest to measure the customer referral value (CRV), defined as an individual customer's contribution to the service provider's goals through referral behavior (Helm 2003). At least two features support the use of this metric as a proxy of customer referral activity: (1) It measures value recommendation at the individual level, which facilitates firm decisions about whom to target with marketing actions; and (2) it can be compared with other behavioral and financial metrics, such as customer lifetime value.

While CRV is relatively easy to define and to understand, it is, on the other hand, not easy to operationalize. In order to model CRV it is necessary first to identify a proxy of referral behavior, second, to design a method for estimating the number of people that the focal customer would bring in to the firm because of his/her recommendation, and finally to define a proxy of customer economic value (i.e. Customer lifetime value) in order to estimate how worthy his/her recommendations are. Surveying the marketing literature most metrics of CRV use either formal models or surveys to measure the process of customer recommendation. For example, Hogan, Lemon, and Libai (2004) propose a customer lifetime value (CLV) model that incorporates customer recommendations as a parameter, such that the greater a person's propensity to recommend, the greater worth he or she has for the company. Using a survey-based method, Kumar, Petersen, and Leone (2007) proposed measuring the referral value of a customer, such that CRV estimates the lifetime value of people who would not have become customers if they had not been referred, as well as the savings in acquisition cost from customers who would have joined the firm anyway, without the referral.

We acknowledge the important contributions of these studies, though formal models and methods based on surveys suffer some shortcomings for the estimation of CRV. First, the traceability problem notes that the purchase of a product rarely can be attributed to a specific referral alone (Helm 2003), and this is the typical assumption of methods based on surveys. Second, formal models ignore how social relationships might affect the likelihood to recommend, even though, depending on the social structure in which the focal customer is embedded, it may be more or less likely that his or her recommendation will have an impact on other actors (Strang and Brandon Tuma 1993). Finally, several scholars questioned the accuracy of surveys on influence research. In part, these inaccuracies arise because recollections of past events suffer from severe memory bias, especially recency bias, and the natural and human tendency of naming people to whom the individual is familiar with rather than the person who made the recommendation (Bernard et al. 1984; Watts 2007)

In summary, CRV provides a good metric of customers' advocacy, because it allows comparisons with other customer-based metrics and can be estimated at the customer level. Existing models of CRV suffer from some flaws in their assumptions and formulation that ultimately affect their calculations. Therefore, this research proposes a new methodology to assess CRV that relies on social network theory, which can accurately capture the essence of social relations and their influence on social contagion.

The remainder of this article is structured as follows: In the next section, we develop a conceptual framework to estimate customer referral value, and then illustrate the proposed model with an empirical application, measuring the CRV for the entire customer base of our empirical sample.

2. CONCEPTUAL FRAMEWORK: A NETWORK MODEL OF CRV

2.1 Why is network theory a good way to model social contagion?

Studies of the phenomenon of recommendation and imitation as a social contagion process mainly use diffusion models (Bass 1969; Mahajan et al. 1984; Van Den Bulte and Lilien 2001) that consider how the adoption of an innovation might be driven by the recommendation of those who have adopted before. Two conflicting assumptions about how social structure and social interaction affect the diffusion of an innovation emerge: If the model assumes spatial and temporal homogeneity, all members of the population have the same chance of affecting and being affected by others, and the potential influence of prior adoption events does not change with the length of time since their occurrence (Strang and Brandon Tuma 1993: 615). Formal models usually assume homogeneity in social relations because the parameter of recommendation is typically modeled as the average of the customer's propensity to recommend (i.e., Hogan et al. 2004) no matter how this individual is connected to whom he/she makes the recommendation to. Thus, assuming homogeneity in social relations is not an accurate representation of the influence of social relationships on people's behavior because a person likely exerts differential efforts to convince a friend or sibling to do something compared with the effort to persuade an acquaintance.

Conversely, the assumption of temporal and spatial heterogeneity implies that actors within a social network have different likelihoods of being influenced during the diffusion process, and this probability might change over time. Academic researches show how relevant is to take into consideration the network structure to understand and to predict social influence. For instance, Frenzen and Nakamoto (1993) and Bone (1992) find that weak ties are not as effective as strong ties for encouraging the flow of certain types of information. Therefore, the structure of social relationships makes a difference in the probability of information transmission and likelihood of social influence.

To capture the heterogeneity of social relations in a diffusion process, previous studies used network models, especially to determine the extent to which an actor's adoption behavior is a function of another actor's knowledge, attitudes, or behavior with regard to an innovation (Burt 1987; Davis and Greve 1997; Strang and Brandon Tuma 1993; Van Den Bulte and Lilien 2001; Van Den Bulte and Wuyts 2007). These models offer insights into how social influence occurs and provide a greater understanding of how the network structure might enable or deter this process. We therefore model the referral activity of a focal customer using a social network approach.

2.2 Definition of Referral Activity: We define the referral activity of customer *j* with respect to a prospect or customer *i*, as a process that takes place either by recommendation, imitation or social pressure, by which *i* might end up adopting or buying a product due to *j's* social influence.

Notice that in the aforementioned definition our concern is not about the causes of social influence (i.e. recommendation, imitation, etc.), but whether this social contagion process (no matter how it takes place) is a significant driver of adoption. We do not focus on the drivers of social influence because in most of the cases it is hard to isolate the effect of each driver. For example a person might likely adopt a new technology because someone told her/him about it, or because he observed what others did with such product. Thus identifying exactly the reasons by which this person bought the device is extremely difficult.

2.3 Modeling Customer Referral Value

We model the CRV of customer *j* as the sum of the economic value generated by the *n* customers *i's* within the social network of *j*, who adopted after him/her, weighted by the probability that each *i* becomes a customer due to *j's* social influence. Formally,

$$CRV_j = \sum_{i=1}^{n} P(CR_{j/i}) \times CV_i \,, \tag{1}$$

where CV_i is the economic value generated by customer *i* because of his or her own consumption. We suggest estimating this metric using a customer lifetime value model (CLV), because there is strong evidence about the benefit of using CLV to assess the customer's future economic value (Berger and Nasr 1998; Dwyer 1997; Reinartz et al. 2005; Rust et al. 2004).

Furthermore, $P(CR_{j/i})$ captures the probability of adoption by customer *i* due to the social influence of customer *j*. We model this probability in two steps. First, we estimate how the focal customer (node *j*) contributes to the personal network exposure (PNE) of node *i*. Second, we estimate how likely node *i* is to become a customer, given his or her personal network exposure to node *j*. We explain the complete process bellow.

<u>Estimating personal network exposure:</u> We assume that social influence stems from exposure to persons who have already adopted and operates on the basis of personal relationships. Thus, the social influence to which person *i* is subject at time *t* is a function of whether other people *j* have adopted previously (indicated by $Y_{j,t-1,}$ which is a vector of ones and zeros of size n x 1)and how important each node *j* is to *i* (indicated by the social matrix $W_{i,j}$ of size n x n). Following Valente (1995), we rescale this metric in terms of proportions by dividing the product of the social weight ($W_{i,j}$) times the vector of adoption $(Y_{j,t-1)}$) by the sum of the entire set of relations of node *i* (indicated by $\sum W_i$ of size n x 1):

$$PNE_{i,t} = \frac{\sum W_{ij} Y_{j,t-1}}{\sum W_i} \,. \tag{2}$$

Equation 2 provides the PNE of node *i* with respect to node *j* and shows the proportion of actors within the network of node *i* that adopts before time *t*. For example, in Figure 1a, we provide the PNE of node *D*, assuming a dichotomous relationship between *D* and his or her contacts. Because node *D* has five contacts and three of them already have adopted, the PNE of node D is 60% at this time. It is worth noting that PNE might vary if the tie strength changes. A strong tie represents a close relationship like a family tie or friendships. These ties are characterized by high levels of trust and exchange of information and knowledge. Thus in figure 1a if the link between D and A would be 5 instead of 1, PNE of node D at this time would be 78%.

----------------------------------Insert Figure 1a about here---

Research in sociology uses PNE as a proxy for social contagion (Burt 1987; Strang and Brandon Tuma 1993; Valente 1995; Van Den Bulte and Lilien 2001), and we believe it offers a good metric for modeling social influence because it captures not only the normative and social pressures of a set of people on the adoption behavior of an individual but also the collective process of social influence, such that a single person's decision can be affected by choices made by many other people, not just one individual. This feature is important; one of the main criticisms of formal models of referral behavior has been their faulty assumption that people make decisions based solely on a single recommendation source.

To assess individual referral activity using PNE, we capture the contribution of node *j* to the personal exposure of node *i* when the latter adopts. For example, in Figure 1a, the contribution of node *F* to the personal exposure of node *D* is 20%, because the strength of ties is equal for all connected nodes. This value captures the specific influence of node *j* on node *i*'s PNE, and therefore, it can be understood as a proxy of the referral activity of node *j* with respect to actor *i*.

Because node j can connect to many other actors in the social network, his or her referral activity should be measured by taking into consideration his or her influence on the personal network exposure of those nodes that they later adopt. Figure 1b depicts this situation: Node F influenced the adoption process of nodes D and G, with contributions to their PNE of 20% and 33%, respectively. If this figure represents the entire set of relationships among these actors, the referral activity of node F can be estimated according to its contribution to the PNE of nodes D and G.

--------------------------------------Insert Figure 1b about here--

Validating the role of personal network exposure in the adoption process: Although PNE might be an important driver of the decision to adopt or buy a product, other factors might affect such behavior as well (e.g., advertising, prices). Thus, we statistically assess the significance of PNE as a driver of adoption and therefore estimate a discrete time logit model following Allison´s methodology (1982), which includes PNE as a predictor, while controlling for other factors that might drive adoption as well. Formally:

$$\log\left(\frac{\Pr(y_{i,t}=1)}{1-\Pr(y_{i,t}=1)}\right)=\beta_0+\beta_1 PNE_{i,t}+\beta_2 X_{i,t}+\beta_3 K_i+\beta_4 J_t\,, \tag{3}$$

where $y_{i,t}$ is a binary vector of adoption behavior, $\mathrm{PNE}_{(i,t)}$ is the personal exposure of customer i at time t, $X_{(i,t)}$ is a set of time-variant covariates that affect the process of adoption, K_i refers to individual-level variables that might affect the customer's adoption, and J_t is a set of time-dependent variables that might impact a customer's decision to adopt.

Notice that another advantage of PNE as a proxy of social influence derives from its flexibility. As Equation 2 reveals, the social network matrix $W_{i,j}$ captures the relevance of j to i with regard to the different pattern of relations between both types of nodes. Thus, $W_{i,j}$ might represent two people who are neighbors, relatives, members of the same organization, and so forth. As a result, a researcher using this metric can determine which kind of social interaction best explains the pattern of influence. We show that in our empirical application.

Estimating customer referral value: The results of the discrete time hazard model can indicate whether PNE is a significant predictor of adoption and reveals the parameter estimates of the logit model that assess the probability that a person will adopt, given certain values of Xs. If PNE is a statistically key driver of adoption, the referral value of *j* with respect to *i* equals the product of *i*'s customer value (CV in Equation 1) and the difference between the probability of adoption of customer *i* due to his or her personal network exposure to node *j* and the probability that *i* adopts without any exposure to the influence of node *j* (see Figure 2).

--------------------------------------Insert figure 2 about here---

Formally:

$$P(CR_{j/i}) = \frac{1}{1+e^{-(\beta_0+\beta_1 PNE_{j/i}+\beta_k X_i)}} - \frac{1}{1+e^{-(\beta_0+\beta_k X_i)}}, \tag{4}$$

where $PNE_{j/i}$ is the contribution of node *j* to the personal exposure of node *i* when the latter adopts; β_o, β_1, and β_k are parameter estimates from the results of the discrete time logit model; and X_i is a vector of variables (both time-variant and time-invariant) that affect *i*'s decision to adopt, regardless of influence from others. In turn,

$$CRV_j = \sum_{i=1}^{n}\left(\frac{1}{1+e^{-(\beta_0+\beta_1 PNE_{j/i}+\beta_k X_i)}} - \frac{1}{1+e^{-(\beta_0+\beta_k X_i)}}\right)\times CV_i\,. \tag{5}$$

Therefore, CRV_j equals the sum of the product between the likelihood that the referral activity of *j* affects each *i*'s decision to adopt, multiplied by *i*'s life time values. To compute how *j* contributes to the personal exposure of *i* when this prospect adopts, we propose a simple method to compute *j*'s contribution to *i*'s personal exposure in the Appendix.

3. EMPIRICAL VALIDATION

3.1. Data and PNE estimation

We test our methodology using a database of Internet adoption in a small town in Argentina. Given the richness of our data we next provide a detailed explanation of our sample.

We obtain the information from a small phone company which is the unique telecom provider of fixed phone service in the above town. The firm granted us access to a database of all telephone calls from 1998 to 2007. During this period, the company had 1,025 residential customers, which include residential households with fixed phone numbers but not small businesses or professional offices. For each year we have data about who each subscriber calls, which can be either to another individual living in the town (in this cases is a local phone) or to another person outside the town (regional, cellular or international call). The database provides information of the number of phone calls that a customer did to a specific subscriber, and about the amount of minutes of each conversation. It is worth noting that during the time period of analysis some people became customers or left the company while others remained with the phone company over the entire study period. In addition, the firm does not charge for local call phones, in other words, people can make unlimited calls to others within the town without having to pay for them.

Of the overall customers, 249 also subscribed to Internet access sold by the company. Because our approach for estimating CRV requires an understanding of referral activity as a diffusion process, we focus on the Internet service, which the company began to sell in 1998, so we have complete adoption and defection data for each customer. Two kinds of technologies for Internet access were sold by this firm between 1998 and 2007: Dial up and DSL. Given the delay in the diffusion of technology in small towns in Argentina, broadband connection through ADSL began to be sold in 2003, at least three years later than in big cities like Buenos Aires, Rosario or Cordoba, whereas dial up access was the first Internet access offer and still some people used it during the period of analysis.

The phone calls made by the 1,025 customers during 1998–2007 were used as a proxy of the structure of social relations among customers. On the basis of this information, we created year-based socio-matrixes that take into account the number of incoming and outgoing calls. This proxy captures the intensity of social interactions among members of the community, which makes it a good benchmark of information transfer processes, as a recent study shows that a high percentage of people make recommendations to friends and colleagues over the telephone (Keller-Fay 2008)

We estimated PNE (Equation 2) for each of the 1,025 customers, by year, taking into consideration the vector of adoption of Internet services at time *t–1* and the socio-matrix of phone calls at time *t*.[1] In this case, PNE (Equation 3) reveals the influence on a focal node from her colleagues through social recommendations or communications; we refer to this variable as WOM (for word-of-mouth communication).

Using a similar approach, we determine whether the process of adoption relies on what Davis and Greve (1997) and Bell and Song (2007) call the geographic proximity or neighborhood effect, which measures whether exposure to the actions of spatially proximate others (by either communication or imitation) influences the trial decisions of those who have yet to experience the service.

To the best of our knowledge, there is not a unique methodology to determine who is a neighbor of who. For example Nam, Manchanda and Chintagunta (2007), following the social interaction literature, define geographically closed neighbors as those who live within a geographic band. Authors choose a circle of 0.5 miles radius centered at any individual such that those people living in this are considered neighbors. Bell and Song (2007:367) in their

[1] For example, when we estimate PNE in 1999, we consider the vector of adopters in 1998 and the socio-matrix that reflects incoming and outgoing calls during 1999.

model consider neighborhood relationships at the zip code level, as they explain: "In our application neighborhood relationships are known at the level of the region (zip code) but not at the level of the individual. We know the exact spatial proximity of different zip codes, but nothing about relative locations of individuals residing in the same zip code. Moreover, no individual-level covariate information is available for either triers or non-triers". Given that existing literature proposed modeling neighborhood relationships taking into consideration geographical proximity, considering the characteristics of our data, we model neighbors' relationships using the following approach: a) We consider neighbors those customers living on the same block (both sides of the street – figure 3a), b) For customers living on the corners, neighbors are those who live on the corners surrounding the same streets, or live within 50 meters of this reference point (figure 3b).[2]

----------------------------------Insert figure 3a and 3b about here------------------------------------

Unlike the phone calls socio-matrix, the proximity matrix is unique, and the edges or links among nodes are dichotomous, such that the only information provided is whether two nodes are neighbors. We again estimate, by year, the PNE of each node, taking into account the proximity matrix and the vector of adoption of Internet services. The PNE in this case therefore indicates how the focal node's adoption behavior is influenced by spatially proximate others. Because nearness is the primary characteristic of this pattern of social relationships, we call this variable Proximity, according to the socio-matrix of neighbors. In summary, we measure social contagion with two proxies: WOM and Proximity.

3.2. Statistical significance of PNE

To determine whether WOM and Proximity are significant drivers of adoption, we estimate the following discrete time logit model:

[2] We confirm the reliability of this methodology in-situ by selecting a random sample and analyzing their connections, according to our methodology. Our classification is reasonably close to people's own perceptions of who their neighbors are.

$$\log\left(\frac{\Pr(y_{i,t}=1)}{1-\Pr(y_{i,t}=1)}\right)=\beta_0+\beta_1 WOM_{i,t}+\beta_2 PROX_{i,t}\beta_3 X_{i,t}+\beta_4 K_i+\beta_5 J_t\,, \qquad (6)$$

where $WOM_{i,t}$ is the PNE of customer i at time t according to the socio-matrix of phone calls, and $PROX_{i,t}$ is the PNE of customer i at time t according to the social relationships among neighbors. In addition, $X_{i,t}$ is a vector of covariates that includes phone consumption and the percentage of out-of-town calls that node i made at time t. We adopt this variable because there might be a substitution effect between phone and Internet usage. Furthermore, K_i is the set of variables at the individual level that might affect adoption, including occupations, because adoption might depend on whether a person works as a farmer, retailer, or professional. Finally, J_t refers to the firm's advertising. However, because investments in marketing did not vary significantly over time, we capture the advertising effect with a dummy variable that indicates the years in which the firm advertised. It is worth noting that this is happened because the firm started advertising the Internet Service in 2003 splitting its tinny fix annual budget between graphic press and radio. As a result, marketing efforts neither were significant nor vary significantly over the time period of the study.

We estimate the discrete-time logit model using maximum likelihood. Significant unobserved heterogeneity occurs in the four models; Rho (quotient of unobserved individual component and the idiosyncratic + individual error) is significantly different from 0. Therefore, we control for intraclass correlation by estimating random effect logit models.

--Insert Tables 1 and 2 about here-------------------------------

In Table 1, we provide the correlation matrix for the independent variables, and in Table 2, we present the results of the discrete time logit model. We first investigate the influence of the control variables on adoption behavior (Model 1), and then introduce the main variables of WOM and Proximity, controlled by the other covariates (Model 2 to 4).

Our estimation relies on the parameter estimates of Model 4, which achieves the highest Akaike information criterion.

Our data reveals that WOM is a significant predictor of customers' adoption of Internet services, but Proximity is not. That is, in this context, social contagion occurs through recommendations or communication among a product's previous adopters and potential ones. Adoption cannot be explained by a neighborhood effect (proximity using our terminology), which contrasts with Bell and Song's (2007) argument and reveals that nearness might not drive adoption behavior. It is worth noting that this result brings evidence about the flexibility and robustness of our model. As we showed before, we model social influence using two proxies: proximity (neighborhood effect) and recommendation (oral communications through phone calls), but only one of the two was found as a statistically significant driver of adoption. As a consequence our approach allows modeling social influence taking into consideration different kinds of social relationships (flexibility) and it also permits to estimate which of these proxies are significant antecedents of adoption (robustness).

3.3. Estimating CRV

Probability that *j* affects *i's* adoption: We estimate the probability that *j* affects *i*'s decision to adopt following the procedure described in the Appendix. We first create a referral weighted matrix (step 2 in the Appendix), in which each link (*i*, *j*) exemplifies the contribution of node *j* to the personal network exposure of node *i* in the year that the latter adopted.[3] We then use Equation 4 to estimate the probability that customer *j* affects the adoption by customer *i*, taking into consideration the parameter estimates of the logit model 4

[3] We followed this approach because of the flexibility of our data. We could create multiple socio-matrixes of incoming and outgoing calls, one per year, to reflect social relationships, but in most situations, researchers likely have access to only one socio-matrix. Therefore, the referral weighted matrix also can be created simply; the link (*i,j*) should reflect how *j* contributes to the personal network exposure of *i* according to the relationship described in the original socio-matrix.

(Table 2), the contribution of *j* to the personal exposure of *i* from the referral weighted matrix, and other individual covariates that might influence *i*'s adoption behavior (e.g., occupation, exposure to advertising). Because the estimation of Equation 4 occurs at the individual level, the covariates in the parameterization are the original values of the time-invariant variables (i.e., 1 if the person is an expert in the use of Internet and the means of the time-variant variables (phone consumption, percentage of out-of-town calls)

Assessing *i's* economic value: According to Equation 1, the assumptions of how to assess the customer lifetime value (CLV) of prospects influenced by the focal customer *j* determine the CRV of the latter.

CLV can be defined as the present value of all future profits obtained from a customer over the life of his or her relationship with a firm (Gupta and Zeithaml 2005). Marketing scholars have proposed multiple metrics of CLV, in which their assumption in the model specification depends among others, on whether there is a contractual relationship between firms and customers, and how retention rate should be estimated (for a review see Villanueva and Hanssens, 2007 and Kumar 2008). As a result, the characteristic of the sample and the specific features of customers–firm relationships would condition how to choose the most suitable CLV model to apply. Thus researchers interested in applying our CRV model should necessarily find the right approach to calculate CLV according to the specific characteristics of the setting which they are dealing with.

In this paper we apply the shifted-beta geometric (sBG) model proposed by Peter Fader and Bruce Hardie for the estimation of CLV, because its assumptions perfectly fit with our data and the characteristics of our sample. Fader and Hardie (2007) introduce the sBG distribution to model customer contract duration; authors showed that this approach provides remarkably accurate forecasts and other useful diagnoses of customer retention. Specifically, we estimate the sBG distribution parameters to fit the model to the multi-cohort data following the procedure suggested by the authors (Fader and Hardie 2007), then we estimate the discounted expected residual lifetime (DERL) for each customer cohort (Fader and Hardie 2009). Finally, we estimate CLV at the customer level by multiplying each customer average margin by the DERL that corresponds to his or her cohort.

3.4. Results

We present the CRV results in comparison with the estimation of CLV at the customer level to clarify which customers are most valuable due to their individual consumption and which ones are worthwhile because of their referral activity.

In Figure 3, we depict the relationship between customer lifetime value and customer referral value. In line with Kumar and colleagues (2007), we find a weak correlation between CRV and CLV (Pearson correlation: 0.13).

--------------------------------------Insert Figure 4 about here--------------------------------

As it can be seen in our empirical application, customer referral value is lower than customer lifetime value, which means that it is hard to find customers whose value as recommenders exceeds their individual consumption value. In contrast, Kumar and colleagues (2007) report that the CRV of a group of customers reaches at least three times their CLV. This difference with our results might reflect prior assumptions about referral activity. That is, Kumar and colleagues (2007) assume that a person who becomes a customer after a recommendation makes that decision because of the advice of a single person, whereas we recognize that a customer attracted by recommendations might be influenced by many other customers who have adopted previously.

On average, referral economic value is lower than individual economic value, yet customer recommendation activity is not worthless. According to our estimation, CRV represents 11% of CLV. This ratio comes from our analysis of the mean proportion of CRV over CLV.

In summary, our results indicate that CRV entails a rather low percentage of the customer economic value and that the correlation between the constructs is fairly weak, which implies that the most profitable customers are not necessarily those with greater influence over others' decisions to adopt. Although we do not find any customers whose CRV is notably greater than their economic value, the results suggest that referral activity is not economically worthless.

4. DISCUSSION

That WOM has a pervasive impact on consumers' behavior has been well documented in previous research. As a consequence, firms are interested in identifying customers who engage in more referral activity, which they might use to attract new clients with minimal financial expenditures. In this regard, CRV, or the economic value that a customer is likely to generate through his or her referral activity, provides a straightforward way to segment the customer base according to their likelihood of influencing others.

We propose a methodology for estimating customer referral value, using social network theory in the model design and event history analysis in its statistical validation. Specifically, we assume that social contagion stems from personal network exposure to those who have already adopted a product. After confirming, with an econometric model, that this proxy of social contagion is a significant driver of adoption, we estimate the CRV of customer j as the sum of the product of the contribution of j to the adoption process of customers embedded in his/her network that adopt subsequently and the economic value of each of these actors.

At least three reasons justify our use of a network model. First, previous studies of WOM referral have not accurately captured the process of social contagion, because they ignore connections among individuals. Informal communications, which provide the framework for the process of social recommendation, should be studied as a social phenomenon, because people do not act in isolation but rather from within a social structure that both influences and is influenced by them. Second, social network theory offers a conceptual toolkit that allows for a better understanding of the main drivers of referral activity. For example, it would be interesting to check whether highly connected persons (high in- and out-degree centrality) and those with stronger tie relationships are more likely to obtain a greater CRV. Third, the combination of network models and event history analysis enables us to determine which pattern of social relationship best explains the process of social contagion in innovation diffusion, because social networks can emerge from different patterns of relationships (friendship, advice, neighbors). The combination makes the model more flexible (different social networks) and robust (it can test statistically which pattern of social relationships explains the process of adoption).

In our empirical application, we find a weak correlation between CRV and customer economic value, in line with prior studies (i.e., Kumar et al. 2007), which confirms that firms should segment their customer base according not only to the value that a customer is likely to generate at the individual level but also the economic value of his or her referral activity. Although we find that CRV represents only a small fraction (about 10%) of a customer's overall economic value, this result refers exclusively to people's willingness to recommend. In other words, the firm did not encourage customers to spread the word or promote its products. Therefore, further research should investigate the extent to which information derived from CRV can help the firm design better customer recommendation programs.

We also acknowledge some limitations in our application of the methodology. First, the model requires access to an entire set of social relationships among groups of customers.

Second, the influence of a specific customer might extend beyond his/her direct links, though we do not account for the influence of these indirect links. This influence may be relatively low, estimated as the influence of node j on the PNE of his or her direct tie, times the direct ties' influence on node j's indirect ties. Because PNE falls between 0 and 1, this multiplication should result in a very low value, but it might be important to consider.

Third, our conclusions regarding the drivers of CRV apply only to industries that have contractual relationships with their customers. However, this limitation does not invalidate the broad application of our proposed model to test CRV in other industries.

Table 1. Correlation matrix of discrete time logit model variables

	1	2	3	4	5	6	7	8
1 Adoption								
2 WOM	0.118*							
3 Proximity	0.032*	0.211*						
4 Advertising	0.044*	0.438*	0.285*					
5 Farmer	0.051*	0.092*	0.022	-0.045*				
6 Retailer	0.026*	0.056*	0.024*	-0.033*	-0.121*			
7 Professional	0.086*	0.08*	-0.002	-0.037*	-0.073*	-0.052*		
8 Phone Consumption	0.185*	0.086*	0.016	-0.048*	0.112*	0.142*	0.114*	
9 % Out-town calls	0.112*	-0.121*	-0.138*	-0.397*	0.061*	0.075*	0.053*	0.329*

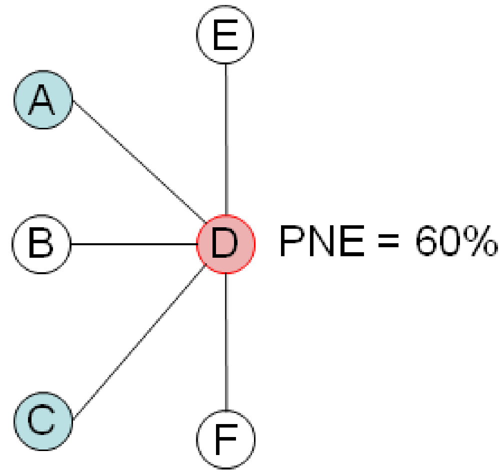
*Significant at 5%.

Table 2. Results of the discrete time binary logit model

	Model 1	**Model 2**	**Model 3**	**Model 4**
WOM		4.10		4.10
		[0.644]**		[0.647]**
Proximity			0.900	0.85
			[0.613]	[0.629]
Advertising	3.32	2.916	3.279	2.87
	[.461]**	[0.465]**	[0.465]**	[0.465]**
Farmer	1.02	0.801	1.006	0.79
	[.448]**	[0.233]**	[0.231]**	[0.233]**
Retailer	0.447	0.337	0.427	0.32
	[.237]	[0.237]	[0.238]	[0.237]
Professional	2.06	1.746	2.040	1.74
	[.329]**	[0.330]**	[0.323]**	[0.330]**
Phone consumption	0.01	0.002	0.002	0.00
	[0.000]**	[0.000]**	[0.000]**	[0.000]**
% Out-town calls	4.93	4.922	4.965	4.95
	[0.437]**	[0.442]**	[0.438]**	[0.443]**
Constant	-10.83	-10.609	-10.822	-10.61
	[0.591]**	[0.592]**	[0.590]**	[0.592]**
Observations	7612	7612	7612	7612
Number of groups	1026	1026	1026	1026
Obs per group (min):	1	1	1	1
Obs per group (avg):	7.42	7.42	7.42	7.42
Obs per group (max):	10	10	10	10
AIC	1794	1777	1840	**1774**
Log Likelihood:	-891.14	-873.76	-890.11	-870.88
Rho:	0.43	0.17	0.438	0.43
Rho (se)	[0.031]	[0.035]	[0.031]	[0.032]

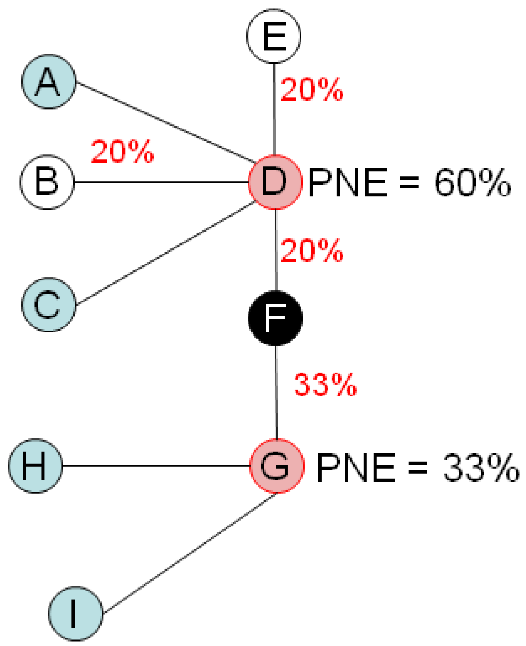
Notes: Dependent variable is adoption at time t. Standard errors in brackets. Coefficients for time dummy variables are not shown.
* Significant at 5%. ** Significant at 1%.

Figure 1a. Personal network exposure of node D



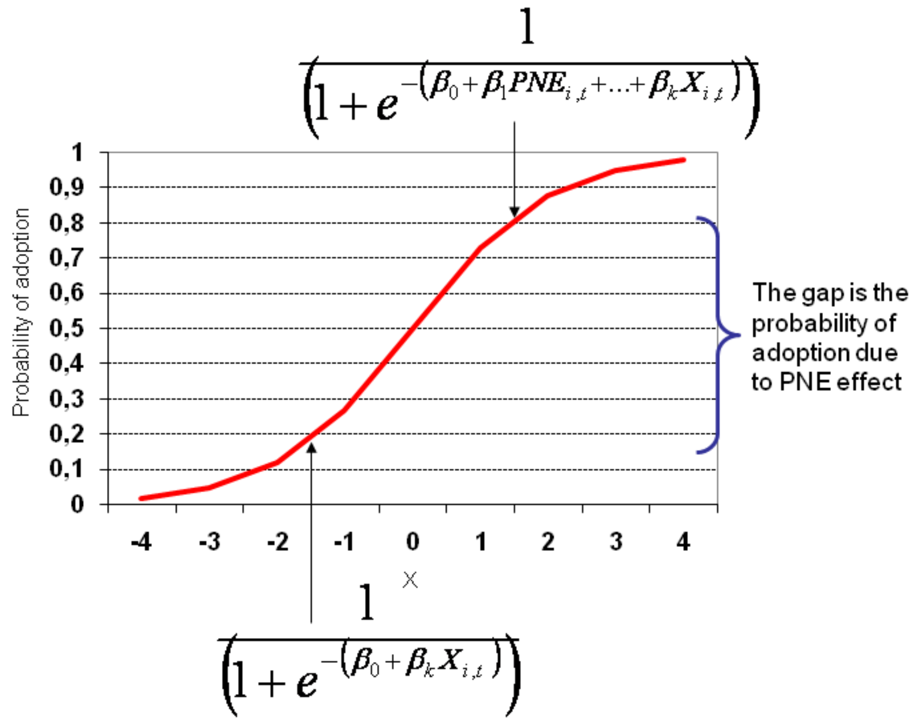
Notes: Nodes F, B, and E adopted before time t, while nodes A and C have not yet adopted.

Figure 1b. Referral activity of node F



Notes: Nodes G and D adopted at time t. Nodes F, B, and E adopted before time t. Nodes A, C, H, and I have not yet adopted.

Figure 2. Estimating the probability of adoption

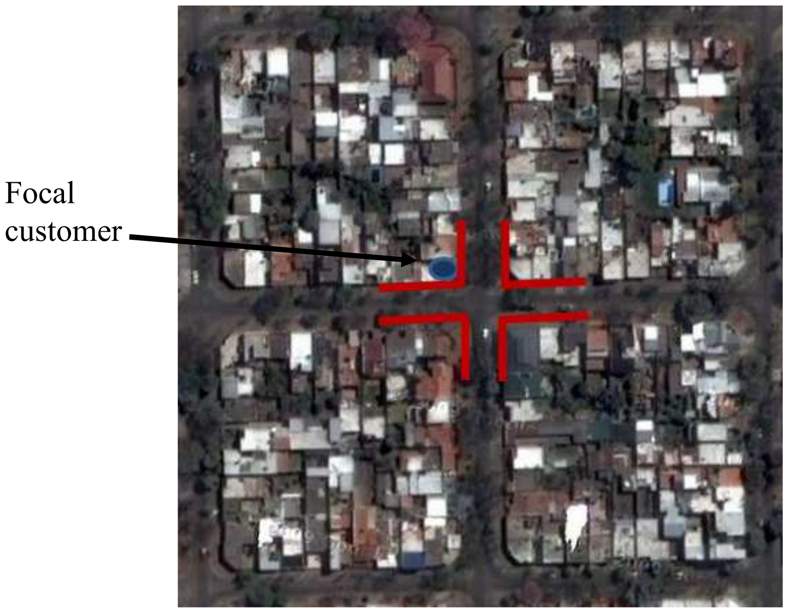


To be read as the probability of adoption due to PNE effect is the difference between the probabilities of adoption of customer *i* due to his or her personal network exposure to node *j* and the probability that *i* adopts without any exposure to the influence of node *j*

Figure 3a. Neighbors of customers living on the same block

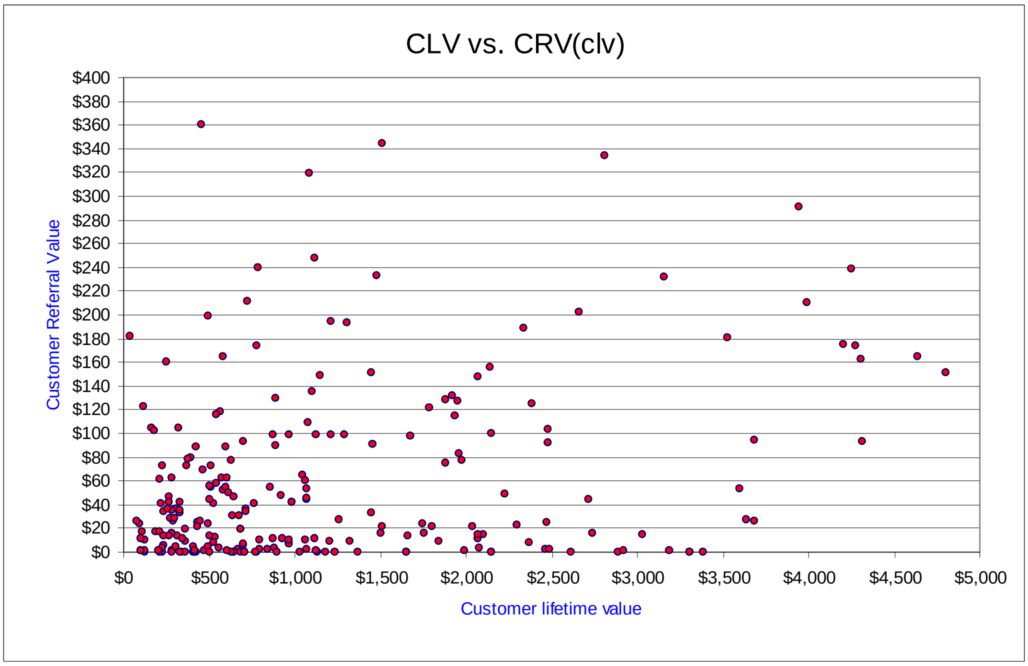
To be read as: Neighbors of the focal customer living where blue circle is located, are those who live in the same block of the former.

Figure 3b. Neighbors of customers living on the corners



To be read as: Neighbors of the focal customer living where blue circle is located, are those who live on the corners surrounding the same streets, or live within 50 meters of this reference point.

Figure 4. Relationship between customer referral value and customer lifetime value



Pearson correlation: 0.13.

Appendix. Procedure to estimate the contribution of *j* to the personal exposure of *i*

The analysis of the contribution of node j to the personal exposure of node i, when the latter adopts, consists of a series of steps.

Step 1. In the analysis of the weighted matrix, because PNE, from Equation 2, is the proportion of influence that node *j* exerts on the adopter *i*, it is important to understand the connection between PNE and the weighted matrix. Assume $W_{i,j}$ is a socio-matrix that reveals the relationship among four actors, 1, 2, 3, and 4, listed in this order in the matrix. It is worth noting that links among actors might take values different from zero or one (it might be for example 5). Also assume that at time *t–1*, nodes 2 and 3 adopt, and at time *t*, nodes 1 and 4 adopt. Following Equation 2, we can derive the PNE for each node (matrix 4).

If we divide the entire social matrix by the total number of relationships of each node i (matrix 5), we get the proportion of influence of each node j in the set of relationships of a focal node i. The weighted matrix then shows the contribution of each node to the PNE of the others. Because nodes 2 and 3 contribute 33.3% each in the total set of relationships of node 4 and adopted before 4, for example, the PNE of 4 is 67%. Node 1 contributes to the remaining 33% of the relationships of node 4, but it should not be included in the referral analysis, because it adopted at time *t*. We are interested only in those nodes that contribute to the PNE of the focal node, so we require a method for cleaning the weighted matrix.

$$W_{i,j}\ (1) \begin{bmatrix} 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{bmatrix} \quad A_{t-1}\ (2) \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \end{bmatrix} \quad \sum_{j=1}^{N} i_j\ (3) \begin{bmatrix} 1 \\ 2 \\ 2 \\ 3 \end{bmatrix} \quad \text{PNE}\ (4) \begin{bmatrix} \mathbf{0\%} \\ 50\% \\ 50\% \\ \mathbf{67\%} \end{bmatrix}$$

$$W_{i,j} \Big/ \sum_{j=1}^{N} i_j \ (5) \quad \begin{bmatrix} 0\% & 0\% & 0\% & 100\% \\ 50\% & 0\% & 50\% & 0\% \\ 0\% & 50\% & 0\% & 50\% \\ 33\% & \mathbf{33\%} & \mathbf{33\%} & 0\% \end{bmatrix}$$

Step 2. We clean the weighted matrix with two steps. First, we set the rows of the weighted matrix of non-adopters and innovators (those that adopt during the first period) to 0. Second, for those nodes not included in the previous category, we multiply, element by element, the transpose of the vector of adoption at time *t–1* by the rows of the weighted matrix of those nodes that adopted at time *t*. In the previous example, we would multiply, element by element, the transpose of A_{t-1} times rows 1 and 4, because nodes 1 and 4 adopted at time *t*. Assuming *t–1* is the first period of adoption, we should set rows 2 and 3 to 0. The result of this algebraic operation is a new matrix that we call the referral weighted matrix:

Referral weighted matrix (6)

$$\begin{bmatrix} 0\% & 0\% & 0\% & 0\% \\ 0\% & 0\% & 0\% & 0\% \\ 0\% & 0\% & 0\% & 0\% \\ 0\% & \mathbf{33\%} & \mathbf{33\%} & 0\% \end{bmatrix}$$

In this matrix, the values in each column indicate the contribution of node j to the personal exposure of each node i. For example, in cells (4,2) of matrix 6, 33% indicates the contribution of node 2 to the personal network exposure of node 4 when the latter adopts.

Step 3. To estimate CRV, the process finishes by applying Equations 4 and 5 to every column of the cleaned weighted matrix. Note 1 does not have a CRV, because it adopted during the last period (*t*). Conversely, the CRV of nodes 2 and 3 can be estimated by adding 33% in the variable PNE of Equation 4, controlling for the covariates that affect the adoption of node 4. The CRV is the probability of adoption due to PNE multiplied by the economic value of node 4.

References:

Allison, Paul D. (1982), Event History Analysis - Regression for longitudinal event data. IOWA: SAGE.

Bass, Frank M. (1969), "A new product growth for model consumer durables," Management Science, 15 (5), 215-27.

Bell, David and Sangyoung Song (2007), "Neighborhood effects and trial on the internet: Evidence from online grocery retailing," Quantitative Marketing and Economics, 5 (4), 361-400.

Berger, Paul D. and Nadal Nasr (1998), "Customer lifetime value: Marketing Models and Applications," Journal of Interactive Marketing, Winter98, Vol. 12 (1), 14.

Bernard, H. R., P. Killworth, D. Kronenfeld, and L. Sailer (1984), "The problem of informant accuracy: The validity of retrospective data," Annual Review of Anthropology, 13 (1), 495-517.

Brown, Jacqueline Johnson and Peter H. Reingen (1987), "Social Ties and Word-of-Mouth Referral Behavior," Journal of Consumer Research, 14 (3), 350.

Burt, Ronald (1987), "Social Contagion and Innovation: Cohesion Versus Structural Equivalence " The American Journal of Sociology, 92 (6).

Corey, Lawrence (1971), "People who claim to be opinion leaders: Identifying their characteristics by self-report " Journal of Marketing, 35 (October), 48-53.

Davis, Gerald F. and Henrich R. Greve (1997), "Corporate Elite Networks and Governance Changes in the 1980s," American Journal of Sociology, 103 (1), 1.

Dwyer, Robert F. (1997), " Customer Lifetime Valuation to Support Marketing Decision Making," Journal of Direct Marketing, 11 (4).

Engel, James, Robert Kegerris, and Roger Blackwell (1969), "Word-of-Mouth communication by the innovator " Journal of Marketing, 33 (July), 15-19.

Fader, Peter S. and Bruce G. S. Hardie (2007), "Computing Derl for the sBG model using excel," Technical Note, 017.

---- (2009), "Customer Base Valuation in Contractual Setting: The perils of ignoring heterogeneity," Marketing Science, (Forthcoming).

Frenzen, Jonathan and Kent Nakamoto (1993), "Structure, Cooperation, and the Flow of Market Information," Journal of Consumer Research, 20 (1), 14.

Godes, David, Dina Mayzlin, Yubo Chen, Sanjiv Das, Chrysanthos Dellarocas, Bruce Pfeiffer, Barak Libai, Subrata Sen, Mengze Shi, and Peeter Verlegh (2005), "The Firm's Management of Social Interactions," Marketing Letters, 16 (3/4), 415-28.

Gupta, Sunil and Valarie Zeithaml (2005), "Customer Metrics: The past, the present, and the future in Academia and Practice," Marketing Science Institute, Special Report 05-200, 36.

Helm, Sabrina (2003), "Calculating the value of customers' referrals," Managing Service Quality, 13 (2), 124-33.

Hogan, John E., Katherine N. Lemon, and Barak Libai (2004), "Quantifying the Ripple: Word-of-Mouth and Advertising Effectiveness," Journal of Advertising Research, 44 (3), 271-80.

Iyengar, Raghuram, Christophe Van Den Bulte, and Thomas Valente (2009), "Opinion leadership and social contagion in new product diffusion," MSI Reports, 08-004.

Keller-Fay, Group (2008), "Talk Track survey - Online vs offline word of mouth."

Kumar, V (2008), "Customer Lifetime Value - The Path to Profitability," Foundations and Trends in Marketing, 2 (1), 1 - 96.

Kumar, V, Andrew J Petersen, and Robert P. Leone (2007), "How valuable is word of mouth?," Harvard Business Review (October).

Mahajan, Vijay, Eitan Muller, and Roger A. Kerin (1984), "Introduction strategy for new products with positive and negative word-of-mouth," Management Science, 30 (12), 1389-404.

Murray, Keith B. (1991), "A test of services marketing theory: Consumer information acquisition activities," Journal of Marketing, 55 (1), 10.

Nam, S., P. Manchanda, and P. K. Chintagunta (2007), "The Effects of Service Quality and Word of Mouth on Customer Acquisition, Retention and Usage," Ann Arbor, 1001, 48109.

Reinartz, Werner J., V. Kumar, and Jacquelyn S. Thomas (2005), "Balancing Acquisition and Retention resources to maximize customer profitability," Journal of Marketing, vol. 69, 13.

Reingen, Peter H. and Jerome B. Kernan (1986), "Analysis of Referral Networks in Marketing: Methods and Illustration," Journal of Marketing Research (JMR), 23 (4), 370-78.

Rust, Roland, Katherine Lemon, and Valarie Zeithaml (2004), "Return on Marketing: Using Customer Equity to Focus Marketing Strategy," Journal of Marketing, 68 (January 2004), 18.

Ryu, Gangseog and Lawrence F. Feick (2007), "A penny for your thoughts: referral reward programs and referral likelihood," Journal of Marketing, 71 (January), 84-94.

Strang, David and Nancy Brandon Tuma (1993), "Spatial and Temporal Heterogeneity in Diffusion," American Journal of Sociology, 99 (3), 614-39.

Summers, John (1970), "The Identity of Women's Clothing Fashion Opinion Leaders," Journal of Marketing Research, 7 (May), 178-85.

Valente, Thomas (1995), Network models of the diffusion of innovations. New York: Cresskil, NH: Hampton Press.

Valente, Thomas and Rebecca L. Davis (1999), "Accelerating the diffusion of innovations using opinion leaders," ANNALS, 566 November.

Van Den Bulte, Christophe and Gary L. Lilien (2001), "Medical Innovation Revisited: Social Contagion versus Marketing Effort," American Journal of Sociology, 106 (5), 1409-35.

Van Den Bulte, Christophe and Stefan Wuyts (2007), Social Networks and Marketing. Cambridge: Marketing Science Institute.

Villanueva, J. and D. M. Hanssens (2007), "Customer Equity: Measurement, Management and Research Opportunities," Foundations and Trends® in Marketing, 1 (1), 1-95.

Villanueva, Julian, Shitin Yoo, and Dominique M. Hanssens (2008), "The Impact of Marketing-Induced Versus Word-of-Mouth Customer Acquisition on Customer Equity Growth," Journal of Marketing Research, 45 (1), 48-59.

Watts, Duncan J. (2007), "Challenging the influential hypothesis " in Measuring Word of mouth - Current thinitk on research and measurement of word of mouth marketing, Walter Carl, Ed. Vol. 3: WOMMA.