University of Navarra

# Working Paper

WP No 516

August, 2003

## THE IMPACT OF ACQUISITION CHANNELS ON CUSTOMER EQUITY

Julián Villanueva*
Shijin Yoo**
Dominique M. Hanssens****

* Professor of Marketing, IESE
** Ph. D. Student, The Anderson School at UCLA
*** The Bud Knapp Professor of Management, The Anderson School at UCLA

**IESE Business School - Universidad de Navarra**
Avda. Pearson, 21 - 08034 Barcelona. Tel.: (+34) 93 253 42 00 Fax: (+34) 93 253 43 43
Camino del Cerro del Águila, 3 (Ctra. de Castilla, km. 5,180) - 28023 Madrid. Tel.: (+34) 91 357 08 09 Fax: (+34) 91 357 29 13

**THE IMPACT OF ACQUISITION CHANNELS ON CUSTOMER EQUITY**


## Abstract

Customer equity (CE henceforth) is a powerful new paradigm to evaluate the firm's value and to optimally allocate marketing resources. This paper is focused on the relationship between customer acquisition and CE. We attempt to answer the following four questions: 1) How should customer acquisition channels be categorized to make them meaningful to managers and academics?; 2) How do we measure the effects of different acquisition channels on the firm's performance?; 3) How do we disentangle short-run effect and long-run effects?; and 4) How should the manager allocate a limited budget among the acquisition channels so as to maximize customer equity?

We first propose a way of categorizing customer acquisition channels according to their level of contact and intrusiveness. A vector-autoregressive (VAR) model is used to examine the dynamics of acquisition channels and the firm's performance, and an empirical illustration on a surviving Internet company will be provided. The results show that each cohort (i.e., customers from different acquisition channels) has different short-run and long-run effects on the firm's performance by the subsequent login and purchasing behavior.

Building on previous research on optimal resource allocation, we develop a Marketing Decision Support System (MDSS) to help managers allocate the acquisition budget among different channels with the objective of maximizing customer equity. We illustrate the consequences of naively maximizing the short-term profit and not accounting for differences in the margin contribution of different cohorts.



**Keywords**: customer equity; customer acquisition; VAR; long-run modeling

**THE IMPACT OF ACQUISITION CHANNELS ON CUSTOMER EQUITY**

## Introduction

Customers are valuable assets for the firm, but they can be costly to acquire and to retain. Customers' heterogeneity in the course of their relationship with the firm is reflected in their price sensitivity, lifetime duration, purchase volume, and even word-of-mouth generation. This heterogeneity causes differences in customers' lifetime value (CLV, hereafter), defined as the discounted stream of cash flows generated over the lifetime of a customer. To the extent that different acquisition strategies will bring different "qualities" of customers, the acquisition effort will have an important influence on the long-term profitability of the firm[1]. Indeed, both practitioners and scholars have emphasized that firms should not spend to acquire just any customer, but the "right" kind of customer (Reichheld 1993; Blattberg and Deighton 1996; Hansotia and Wang 1997; Blattberg, Getz, and Thomas 2001). Therefore, the customer acquisition process plays an important role in the newly-emerging paradigm of customer equity (CE)[2].

Optimizing the acquisition budget for long-term profitability is particularly relevant for start-ups and for firms competing in growth markets, where acquisition spending is the most important expense in the marketing budget. In these scenarios, the firm could have an illusion of profitable growth, when in fact it is acquiring unprofitable customers. This occurred for many Internet start-ups that spent aggressively on acquisition in an effort to maximize 'eyeballs', with the hope of locking-in customer revenue later. For many companies, however, that revenue never came, either because their value proposition was not compelling enough or because the underlying linkage between acquisition spending and long-term profitability was poorly understood.

In order to grow their businesses, companies acquire customers using a variety of channels. In this paper, we define an acquisition channel as any vehicle that initially drives a prospect to the firm. While broadcast media and direct marketing are the most traditional acquisition channels, firms also acquire customers through other vehicles such as public

---

1 Moreover, models that do not account for the effect that acquisition has on customer retention will result in biased estimates (Thomas 2001).

2 For a general discussion of the CE concept, see Blattberg, Getz, and Thomas (2001), and Rust, Zeithaml, and Lemon (2000).

NOTE: The authors thank an anonymous collaborating company for providing the data used in this paper. They thank seminar participants at UCLA, UC Riverside, University of Iowa and Dartmouth College for helpful comments.

Research support from the Harold Price Center for Entrepreneurial Studies at UCLA is gratefully acknowledged

relations and word-of-mouth. Thus, it is important to understand the relative effectiveness of these acquisition channels.

In a recent survey of marketing managers of Internet firms, it was found that managers do not predominantly use the channels that they believe are the most effective (Forrester Research 2001). For example, in this study *affiliate programs* was said to be a very effective channel, but it was rarely used. This suggests that managers are unclear about the effectiveness of different channels of acquisition. For example, online ad banners have been criticized as ineffective since they drive few click-throughs and exhibit small conversion rates. However, some authors have warned that media that appear ineffective in the short run may generate consumer awareness and become effective in the long run (Briggs and Hollis 1997; Drèze and Zufryden 1998). Consequently, acquisition-channel effectiveness should be measured with models that can quantify short-run as well as long-run response to marketing stimuli.

The distinction between short and long-run effects is not new in the marketing literature, and several statistical models or experiments capable of capturing this distinction have been proposed[3]. Nevertheless, managers are often criticized as myopic when making spending decisions in that they tend to maximize the short-term and neglect the long-term profitability of the firm. This may occur because the managers' incentives are linked to short-term metrics such as market-share movements. On other occasions managers lack the necessary tools to measure the long-run effects of their decisions. The inability to measure the future consequences of current decisions increases the uncertainty of future payoffs, especially in turbulent markets that are difficult to forecast. By contrast, short-run metrics such as current market share have a strong credibility at all levels of management and are easy to justify (Keil, Reibstein, and Wittink 2001). Nonetheless, neglecting the long-term effects of current actions can lead to suboptimal spending decisions, resulting in inferior long-run profitability and shareholder value creation (Doyle 2000).

Hence, there is an urgent need to develop models capable of measuring the long-run effects of different acquisition strategies, and provide systems to help managers optimally allocate their acquisition spending among different channels. These models should be able to disentangle the long-run from the short-run effects, incorporate the risk associated with future payoffs, and take into account the costs associated with different acquisition channels. This is the main objective of the current paper. Moreover, we depart from "soft" metrics of communication effectiveness (e.g., brand awareness) to "hard" metrics of profitability (Greyser and Root 1999), in that we measure the effectiveness of each acquisition channel with respect to its contribution to the CE of the firm[4]. Once these long-run effects have been measured, we can optimally allocate the acquisition budget among the various channels. In doing so, we do not measure the expected CLV of a customer, but rather her *CE contribution*. In this way a customer is worth not only her own expected CLV but also all the indirect impacts that she has on the firm's performance over time (e.g., by bringing new customers to the firm through word-of-mouth).

---

[3] For example, streams of research include the use of multivariate time-series techniques (e.g., Dekimpe and Hanssens 1995a, 1995b, 1999; Bronnenberg, Mahajan, and Vanhonacker 2000; Pauwels, Hanssens, and Siddarth 2002; Nijs et al. 2001), varying-parameters approaches (e.g., Mela, Gupta, and Lehman 1997; Jedidi, Mela, and Gupta 1999), and experiments (e.g., Lodish et al. 1995; Anderson and Simester 2001). For a review of long-run marketing modeling, see Dekimpe and Hanssens (2000). This paper fits into the emerging literature of linking marketing spending to long-run shareholder value (e.g., Pauwels et al. 2003).

[4] Though there exist various definitions of customer equity (CE), it is defined in this paper as the sum of all existing and expected customers' CLVs. Here CE is used as a metric to show the long-run performance of the firm.

The paper is organized as follows. First, we categorize customer acquisition channels to investigate their short-run and long-run differences with respect to the impact on CE. Second, we propose a VAR model to estimate the long-run effect of a customer acquired from each channel on the long-term performance of the company. Third, we provide an empirical illustration using data from an Internet start-up. Lastly, we develop a marketing decision support system to help the manager allocate her acquisition budget among the different channels.

## Research Development

Our study differs from previous literature on media selection in at least three ways. First, we consider important yet under-researched acquisition channels, such as word-of-mouth and public relations. Second, we study long-run effects of the different acquisition channels on the firm's performance, as separate from short-run effects. Third, we specify the long-run effects of acquisition in a customer-equity framework. Thus, our research is particularly relevant for relationship businesses in which the firm spends aggressively on acquiring customers, in the hope of deriving a substantial future revenue stream. Examples include the wireless telephone industry, broadband Internet service providers, and cable television.

### *Classification of Customer Acquisition Channels*

Our focus on customer acquisition includes all possible channels that drive new customers to a firm, including those that are difficult to control, such as word-of-mouth. An increasing number of firms uses such channels. For instance, BMG Music Service not only spends on online ad banners and direct mail, but also gives referral incentives (in the form of free CDs) to existing customers. Netflix, an online DVD rental firm, spends on online ad banners, places free trial coupons in the DVD-player cartons of some manufacturers, mails other free-trial coupons to targeted audiences and encourages referrals, although without a monetary incentive.

Our classification is based on two dimensions used in previous research. The first is the acquisition channel's *level of contact* with the prospect, which can be personal or broadcast. For example, if customers learn about the firm from a friend or from an email, the contact is more personalized than if they hear about it from a TV advertisement or a newspaper article. Indeed, in the former case someone decided to send a message to that specific customer, while in the latter the message is available to anyone exposed to the medium. Similar to the concept of audience addressability (Blattberg and Deighton 1991), we expect personal contacts to have high addressability and broadcast contacts to have lower addressability[5].

The second dimension is the *level of intrusiveness* of the acquisition channel, which can be low or high. Following the persuasion knowledge model (Friestad and Wright 1994, 1995), we predict that perceived intrusiveness has an impact on customer response and

---

5 Audience addressability is defined by Blattberg and Deighton (1991) as the medium's ability to reach a defined segment of consumers and minimize exposure to other unwanted audience groups. Note, however that level of contact and audience addressability are not exactly the same. For example, a TV ad could have higher audience addressability than a mailing if the TV ad is targeted to a very small and specific consumer segment using a specialized TV program or channel.

subsequent behavior. Indeed, consumers interpret and cope with marketers' communication attempts (e.g., advertising) based on contingent persuasion knowledge. They understand that the main goal of marketing communications is to influence their own beliefs and/or attitudes about the firm's products or services. Thus, we argue that visibly commercial acquisition channels such as direct marketing or mass advertising will be perceived as more intrusive than channels such as public relations or word-of-mouth[6].

Our classification of acquisition channels based on the level of contact and level of intrusiveness results in four categories, namely, word-of-mouth (WOM), direct marketing (DM), advertising (AD), and public relations (PR). A wide array of acquisition tactics can be assigned to one or other of these categories, and we present some of them as an illustration in Figure 1. This classification is managerially relevant, as it includes many non-traditional but widely used acquisition tactics in a comprehensive way. Moreover, it is based on existing consumer behavior theory and therefore we expect these four categories to differ both in their short and in their long-run effectiveness[7].

**Figure 1. Classification of Acquisition Channels**

| | | **Level of Contact** | |
|---|---|---|---|
| | | Personal | Broadcast |
| **Level of Intrusiveness** | High | ***DM***<br>• *Catalogues*<br>• *Email*<br>• *Promotion calls* | ***AD***<br>• *TV ad*<br>• *Radio ad*<br>• *Print ad*<br>• *Online ad banner* |
| | Low | ***WOM***<br>• *Word-of-mouth from friends*<br>• *Referral from search engines* | ***PR***<br>• *Print articles*<br>• *Online articles* |

*Measuring Acquisition Effectiveness*

In this research we develop a metric that helps us link acquisition efforts to shareholder value by measuring the impact of the acquisition spending on customer equity, which has been suggested as a powerful metric for the value of a firm (Gupta, Lehman, and Stuart 2002). Hence, models capable of maximizing customer equity should help managers maximize shareholder value.

---

6 Word-of-mouth communication has been said to be more persuasive than conventional advertising (e.g., Herr, Kardes, and Kim 1991; Brown and Reingen 1987).

7 Nevertheless, the development and testing of formal hypotheses on how level of contact and intrusiveness affect short and long-run effectiveness are beyond the scope of this paper. With our particular dataset (introduced below), we cannot control for personal differences among groups, so we do not know whether those differences are caused by the nature of the medium, or by individual characteristics. We ran a multivariate discriminant analysis of group membership on some personal demographics and found that there are statistically significant differences in the demographics across groups. Therefore, it can be tentatively concluded that different acquisition channels bring different kinds of customers to the firm.

Unlike previous CLV models, our model investigates cross-sectional heterogeneity at the acquisition channel level[8]. For example, previous work has assumed that customers are homogeneous in their expected future value (e.g., Blattberg and Deighton 1996), or longitudinally heterogeneous depending only on the period of acquisition (e.g., Gupta, Lehman, and Stuart 2002). However, we expect different acquisition channels to yield customers that are unequal in their contribution to customer equity. This heterogeneity of acquisition channels has important implications for optimal resource allocation, as firms want to allocate their limited acquisition budget among the different acquisition channels so that they maximize their customer equity and therefore shareholder value. We shall emphasize the differences between the short and the long-run effectiveness to illustrate the importance of maximizing the latter when allocating marketing resources.

## Methodology

*Linkage between Acquisition and Long-Run Performance*

The acquisition process and its link with the firm's performance should be examined as a complex system in which many interactions could take place over time. For example, when computing the marginal contribution of one new customer to CE, we want to measure not only her expected CLV but also all the indirect influences that this acquisition will cause in the firm's performance.

We propose a vector-autoregressive (VAR) model to investigate these interactions which we characterize as follows: (1) *Direct effects* of acquisition on the performance of the firm. We are interested in measuring the impact on the firm's performance (e.g., profits) of a person being acquired from a given acquisition channel; (2) *Cross-effects* among channels. For instance, we are interested in how different acquisition channels affect future word-of-mouth. As an illustration, customers acquired through public relations may generate more referrals than those acquired from direct marketing; (3) *Feedback effects*. The firm's current performance may affect differently the number of customers acquired through different channels in the future. For instance, firms that develop stronger reputations may increase future customer acquisitions through public relations; (4) *Reinforcement effects*. Both firm performance and customer acquisitions will have an effect on each other in the future. For instance, there may be some inertia in the firm that prompts it to use certain channels more and others less if it believes some are more effective than others.

For ease of exposition, assume a three-variable system that captures the dynamic interrelationships among the number of customers acquired at time $t$ through mass advertising ($AD_t$), the number of customers acquired at time $t$ through word-of-mouth ($WOM_t$), and a proxy variable for the firm's performance at time $t$ ($V_t$). The VAR(p) model would be specified as[9],

---

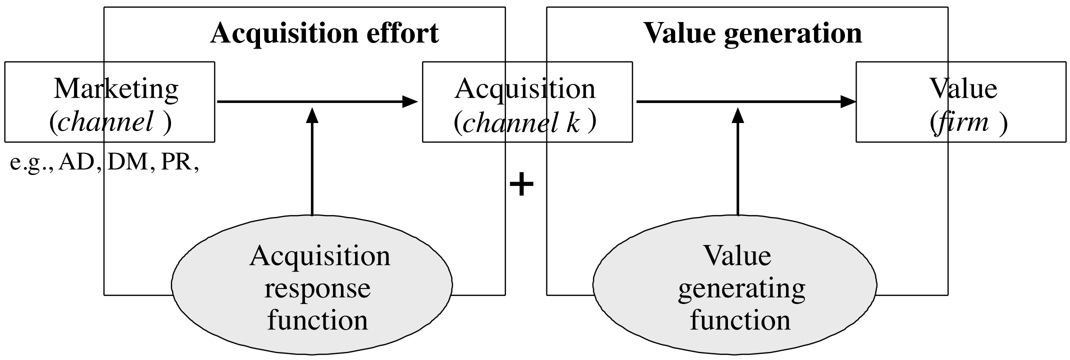
8 We will investigate how each acquisition channel contributes to the firm's customer equity and study heterogeneity for our four categories of acquisition channels. Our measurement approach could nevertheless be used for any particular acquisition channel or for any other categorization. There could also be heterogeneity at different levels, for instance due to demographic characteristics of the individuals attracted by each channel.

9 A deterministic trend, seasonal dummy variables, and exogenous variables can also be included in this VAR. Instantaneous effects are not included directly in this VAR, but they are reflected in the variance-covariance matrix of the residuals ($\Sigma$).

$$(1) \qquad \begin{pmatrix} AD_t \\ WOM_t \\ V_t \end{pmatrix} = \begin{pmatrix} a_{10} \\ a_{20} \\ a_{30} \end{pmatrix} + \sum_{l=1}^{p} \begin{pmatrix} a_{11}^{l} & a_{12}^{l} & a_{13}^{l} \\ a_{21}^{l} & a_{22}^{l} & a_{23}^{l} \\ a_{31}^{l} & a_{32}^{l} & a_{33}^{l} \end{pmatrix} \begin{pmatrix} AD_{t\pm l} \\ WOM_{t\pm l} \\ V_{t\pm l} \end{pmatrix} = \begin{pmatrix} e_{1t} \\ e_{2t} \\ e_{3t} \end{pmatrix}$$

For this VAR model of order $p$, where $(e_{1t}, e_{2t}, e_{3t})'$ are white-noise disturbances following $N(0, \Sigma)$, the direct effects are captured by $a_{31}$, $a_{32}$, cross effects by $a_{12}$, $a_{21}$, feedback effects by $a_{13}$, $a_{23}$ and, finally, reinforcement effects by $a_{11}$, $a_{22}$, $a_{33}$. The researcher could, of course, include additional acquisition channels and even impose restrictions on some of these parameters if there is an a priori reason for doing so. VAR models can be heavily parameterized, depending on the number of variables and time lags in the model. Therefore, long time series are desirable. Note that we do not include marketing activity data (e.g., advertising expenditures, price promotions) since at this point we are not interested in measuring how these marketing efforts lead to number of customers acquired. Instead, we want to measure how much a specific customer contributes to the firm's performance now and in the future. The function linking the number of customers acquired to the contribution to the firm's customer equity will be called the *value generating function*. The interactions between marketing spending and number of acquisitions is captured by an *acquisition response function* (see Figure 2). We will join these two functions later.

**Figure 2. Value Generation through Customer Acquisition**




*Impulse Response Functions and Customer Equity*

Given data availability, a VAR model not only captures all the previous effects (i.e., direct, cross, feedback and reinforcement), it also measures the time dynamics of each effect. We are interested in disentangling the immediate and the long-run effects, and in determining the total cumulative effects. This is accomplished by Impulse Response Functions (IRFs) that trace the present and future response of a variable to an unexpected shock in another variable. VAR models and IRFs have been introduced to the marketing literature in a marketing-mix context (e.g., Dekimpe and Hanssens 1995a, 1995b, 1999; Bronnenberg, Mahajan, and Vanhonacker 2000; Nijs et al. 2001; Srinivasan, Bass, and Popkowski 2000). They are used here to assess how one unexpected customer acquisition, for example from the advertising channel, impacts customer equity over time. To the best of our knowledge, this is the first use of the VAR method to measure the financial contribution of newly acquired customers.

Assuming data stationarity, we can rewrite the VAR model in equation (1) as a moving-average representation (see Enders 1995):

$$(2)\qquad \begin{pmatrix} AD_t \\ WOM_t \\ V_t \end{pmatrix} = \begin{pmatrix} \overline{AD} \\ \overline{WOM} \\ \overline{V} \end{pmatrix} + \sum_{i=0}^{\infty} \begin{pmatrix} \phi_{11}(i) & \phi_{12}(i) & \phi_{13}(\mathrm{i}) \\ \phi_{21}(i) & \phi_{22}(i) & \phi_{23}(\mathrm{i}) \\ \phi_{31}(i) & \phi_{32}(i) & \phi_{33}(\mathrm{i}) \end{pmatrix} \begin{pmatrix} \varepsilon_{1l-i} \\ \varepsilon_{2l-i} \\ \varepsilon_{3l-i} \end{pmatrix}$$

The coefficients $\phi_{\mathrm{jk}}(i)$ are called impact multipliers and measure the impact of a one-unit change in $\varepsilon_{k(t-i)}$ on the $j^{th}$ variable. The different sets of coefficients $\phi_{\mathrm{jk}}(i)$ for $i = \{0,...,\infty\}$ are called impulse response functions and are usually plotted to visualize the dynamic behavior of the variables of interest as a function of shocks in other variables. We can calculate the cumulative long-run effect of unit impulses in any error shock on another variable by accumulating the impact multipliers,

$$(3)\qquad \mathit{Total\ Effect}\,(k \rightarrow j) = \sum_{i=0}^{\infty} \phi_{jk}\,(i)$$

When variables are stationary, the impact multipliers tend to be zero for sufficiently high numbers of $i$ and therefore the total effect is finite[10].

In order to estimate the effect of one new customer acquisition from a specific channel on the long-run performance of the firm we take the following steps: (1) estimate the impulse response functions defined as the effect of a one-person shock in the acquisition channel on the firms' performance ($V_t$); (2) select the impact multipliers that are significantly different from zero; and (3) accumulate significant impact multipliers using a discount rate. Thus, the long-run impact multiplier for a direct effect is obtained as

$$(4)\qquad \gamma_k = \sum_{i=0}^{m} \delta^{\mathrm{i}} \phi_{vk}(i), \delta = \frac{1}{1 + r}$$

where $r$ is the discount rate[11], $m$ is the number of periods to include in the calculation, and $\phi_{vk}(i)$ is the impact multiplier measuring the response of the $V$ variable to the shock of the $k^{\mathrm{th}}$ variable $i$ time units ago.

So long as $V_t$ is a good proxy for the contribution of each customer to the firm's profits, this impact multiplier can be interpreted as the contribution of one customer acquired through a specific channel to the firm's customer equity before accounting for differences in acquisition costs[12]. On other occasions, however, $V_t$ may not be expressed in monetary value. In such cases the impact multiplier needs to be translated to profit contribution, for example,

$$(5)\qquad \lambda_k = \tau(\gamma_k)$$

10 When variables are evolving, the standard procedure is to estimate the VAR model with variables in first differences. In those cases, the IRFs should be accumulated to measure the impact on level forms.

11 The discount rate should incorporate the risk associated with the specific investment. For example, factors such as expected future competition or the urgent need to raise money might affect this rate.

12 Note that $\gamma_k$ does not take into account that a person acquired from a given acquisition channel, say advertising, could be more expensive to acquire than a person acquired through for instance word-of-mouth. We shall come back to this issue later.

where $\tau(\gamma_k)$ is a function that translates the direct effects (as measured by the impact multipliers) on the firm's profits. This approach may be necessary, for example, for an online newspaper that generates revenue from advertising but can only observe its users' login behavior[13]. This login behavior would presumably be highly correlated with advertising exposure and, therefore, with the firm's financial performance.

In conclusion, we have developed a metric that is capable of measuring the long-run CE contribution of a newly acquired customer. This metric captures not only the expected CLV of a new customer, but also all indirect effects that affect the firm's value through that particular customer acquisition.

## Empirical Illustration

*Data Description*

We study an Internet firm that provided free web hosting to registered users during a 70-week long observation period. At the time of registration, individuals provided a demographic profile and responded to the question "How did you hear about our company?"[14], followed by a list of several acquisition channels. Once registered, individuals' unique behavior was tracked as they logged in to use the firm's services (e.g., changing the content or appearance of their web site, or checking on the number of site visits). From these records, we calculate the weekly total number of unique logins, as well as the number of registrations per acquisition channel. These channels are grouped according to the classification in Table 1, where we also show some descriptive statistics. A very small number of registrants who indicated "Other" as their acquisition channel was discarded from this analysis.

**Table 1. Classification of the Acquisition Channels and Descriptive Statistics**

| | Number of Weekly Acquired Customers | | | | |
|---|---|---|---|---|---|
| **Series** | **Mean** | **Median** | **Maximum** | **Minimum** | **Standard Deviation** |
| AD[1] | 305.56 | 317.50 | 550.00 | 56.00 | 116.57 |
| DM[2] | 318.99 | 293.00 | 953.00 | 16.00 | 193.77 |
| PR[3] | 675.01 | 685.50 | 1104.00 | 103.00 | 216.57 |
| WOM[4] | 850.69 | 726.00 | 1746.00 | 67.00 | 420.69 |

[1] AD: Online ad banner; TV, radio, magazine or newspaper advertisement; [2] DM: From an email link; Mailing to your home or business; [3] PR: Mentioned or linked from other web site; Magazine or newspaper article; [4] WOM: Referral from friends or colleague; Referral from professional organization or association; Referral from search engine

[13] This is not necessary if the firm is only interested in finding out which channel is best, assuming that $\tau(\gamma_k)$ is the same for all acquisition channels.

[14] This particular firm did not allow for multiple responses in this question. Therefore, we study the predominant channel that brings a customer to the firm.

The number of logins is a good proxy for the firm's performance given the characteristics of this business[15]. Most free-service Internet companies generate advertising revenue based on logins or click-throughs. Furthermore, once a sufficient number of registrations was achieved, the company switched to a fee-for-service revenue model. As explained in detail in Appendix A, intensity of login behavior was found to have a positive and statistically significant correlation with customers' willingness to pay. Therefore, acquisition channels that yield customers with high usage (login) intensity and therefore a higher probability of converting to a fee-based service, will be considered as the most effective.

The variables are defined as follows:

$AD_t$: number of new registrations at time $t$ from mass advertising
$DM_t$: number of new registrations at time $t$ from direct marketing
$PR_t$: number of new registrations at time $t$ from public relations
$WOM_t$: number of new registrations at time $t$ from word-of-mouth
$V_t$: total number of unique logins at time $t$

*VAR Estimation*

The VAR estimation begins with a unit-root test to determine whether the series is evolving or stationary (see Dekimpe & Hanssens 1995 for a detailed explanation). We use the augmented Dickey-Fuller (ADF) unit root test (e.g., Enders 1995, p. 257), in which the null hypothesis of unit root corresponds to $H_0 : \rho = 0$ in

(6)
$$\Delta y_t = \alpha_0 + \alpha_1 t + \rho y_{t-1} + \sum_{i=1}^{k} \beta_i \Delta y_{t-i} + \varepsilon_t$$

We apply the iterative procedure proposed in Enders (1995, pp. 256-258) to decide whether to include a deterministic trend in the test. The results are shown in Table 3. All variables except AD were found to be trend stationary at a 95% confidence level. Since it has been argued that conventional unit root tests (e.g., ADF) tend to over-accept the null of unit root, we confirmed our results with the KPSS test (Kwiatkowski et al. 1992), which uses the null of stationarity[16]. We found that the two tests disagree for AD and WOM and concluded that all series seem to be stationary at or near the 95% confidence level.

[15] As pointed out in the previous section, our methodology can be implemented with any other proxy variable for the firm performance.
[16] Note that the null hypothesis of the ADF test is that a series has unit root (i.e., evolutionary).

**Table 2. Unit Root Test Results**

| *Series* | *ADF ($H_0$: unit root)* | | | *KPSS ($H_0$: stationary)* | | |
|---|---|---|---|---|---|---|
| | *Stat* | *5%-crit* | *Unit root?* | *Stat* | *5%-crit* | *Unit root?* |
| AD | -3.00 | -3.48 | Yes | 0.06 | 0.15 | No |
| DM | -5.82 | -3.48 | No | 0.15 | 0.15 | No |
| PR | -3.65 | -3.48 | No | 0.07 | 0.15 | No |
| WOM | -4.14 | -3.48 | No | 0.16 | 0.15 | Yes |
| V | -4.34 | -3.48 | No | 0.08 | 0.15 | No |

We proceed to estimate the VAR in level form including all performance variables, a deterministic trend[17] *t* and a dummy variable *d*,

$$(7) \quad \begin{pmatrix} AD_t \\ DM_t \\ PR_t \\ WOM_t \\ V_t \end{pmatrix} = \begin{pmatrix} a_{10} + \alpha_1 t + \gamma_1 d \\ a_{20} + \alpha_2 t + \gamma_2 d \\ a_{30} + \alpha_3 t + \gamma_3 d \\ a_{40} + \alpha_4 t + \gamma_4 d \\ a_{50} + \alpha_5 t + \gamma_5 d \end{pmatrix} + \sum_{l=1}^{p} \begin{pmatrix} a_{11}^l & a_{12}^l & a_{13}^l & a_{11}^l & a_{14}^l & a_{15}^l \\ a_{21}^l & a_{22}^l & a_{23}^l & a_{21}^l & a_{24}^l & a_{25}^l \\ a_{31}^l & a_{32}^l & a_{33}^l & a_{31}^l & a_{34}^l & a_{35}^l \\ a_{41}^l & a_{42}^l & a_{43}^l & a_{41}^l & a_{44}^l & a_{45}^l \\ a_{51}^l & a_{52}^l & a_{53}^l & a_{51}^l & a_{54}^l & a_{55}^l \end{pmatrix} \begin{pmatrix} AD_{t\pm l} \\ DM_{t\pm l} \\ PR_{t\pm l} \\ WOM_{t\pm l} \\ V_t \pm l \end{pmatrix} + \begin{pmatrix} e_1 \\ e_2 \\ e_3 \\ e_4 \\ e_5 \end{pmatrix}$$

The dummy variable is included in order to achieve multivariate normality of the model residuals. This assumption will be needed when deriving the generalized impulse response function (cfr. infra) (Koop, Pesaran, and Potter 1996; Pesaran and Shin 1998). Estimating the model without the dummy variables yields residual outliers in five weeks, After accounting for these outliers, the MVN assumption is met, following the Lutkepohl test (1993, p.155-158). We also test for residual autocorrelation with a portmanteau test (Lutkepohl 1993, p.150-152) and find that the null hypothesis of white noise cannot be rejected.

We find the optimal lag length to be one, using Schwartz's Criterion. Although this VAR model uses 8 parameters per equation, there are sufficient time-series observations (70 weeks) to estimate them. The estimation results are reported in Table 3, and the impulse response functions are shown in Figure 4. We use generalized impulse response functions because imposing a temporal ordering of the variables is not credible in our case. These IRFs are plotted for |t|-statistics exceeding 1, following the procedure in Dekimpe and Hanssens (1999).

[17] The standard practice in VAR modeling is to include a deterministic trend when variables are shown to be trend stationary. The decision of whether or not to include the trend is, nevertheless, not trivial. In particular, Sims (1980) and Doan (1992) argue against detrending because that may discard information concerning the co-movements in the data. We decided to detrend the data, following previous VAR modeling in marketing, and because in this application the trend is likely caused by the natural evolution (growth) of the internet market.

**Table 3. VAR Model Estimation Results**

| | *AD* | | *DM* | | *PR* | | *WOM* | | *V* | |
|---|---|---|---|---|---|---|---|---|---|---|
| AD(-1) | 0.751 | (3.803) | -0.270 | (-0638) | 0.616 | (1.773) | 0.147 | (0.396) | 2.878 | (1.202) |
| DM(-1) | -0.000 | (-0.006) | 0.245 | (1.518) | -0.098 | (-0.739) | -0.428 | (-3.025) | -1.133 | (-1.241) |
| PR(-1) | 0.095 | (0.580) | 0.532 | (1.510) | 0.717 | (2.477) | 0.403 | (1.303) | 1.306 | (0.655) |
| WOM(-1) | -0.024 | (-0.240) | -0.354 | (-1.647) | -0.077 | (-0.434) | 0.594 | (3.146) | 0.727 | (0.598) |
| V(-1) | -0.014 | (-0.625) | -0.019 | (-0.408) | -0.056 | (-1.434) | -0.039 | (-0.937) | 0.124 | (0.460) |
| Intercept | 53.672 | (1.550) | 66.083 | (0.891) | 165.248 | (2.711) | 61.178 | (0.939) | 1,021.253 | (2.432) |
| Trend | 2.982 | (1.052) | 10.051 | (1.655) | 12.278 | (2.461) | 13.550 | (2.540) | 139.216 | (4.051) |
| Dummy | -27.091 | (-0.814) | 226.508 | (3.179) | 37.216 | (0.635) | -131.915 | (-2.108) | -301.534 | (-0.747) |
| R Squared | 0.637 | | 0.395 | | 0.667 | | 0.899 | | 0.953 | |
| F Statistic | 15.297 | | 5.693 | | 17.432 | | 77.725 | | 178.486 | |

*Results*

We interpret the direct, cross and feedback effects of customer acquisition shocks. These are the most insightful managerially, in particular the direct effects, as they will determine the shape of the value generating function.

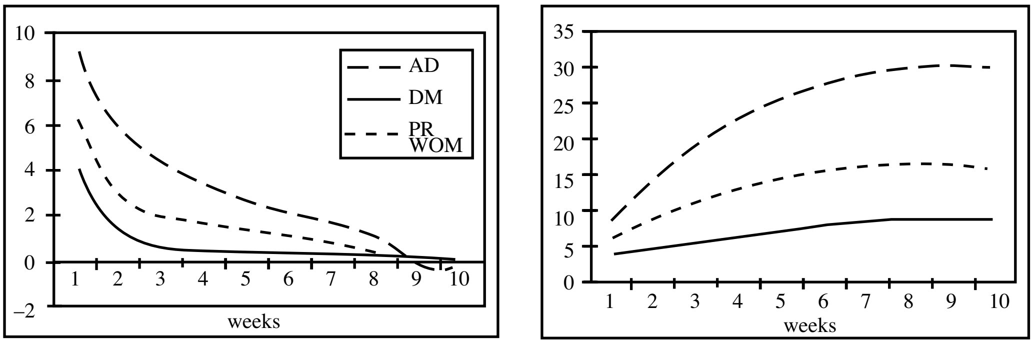
*Direct Effects*. These IRFs measure the total or net effect of an unexpected acquisition on the firm's performance, defined as the total number of logins over time. The net effect includes not only a new customer's own login activity, but also the effect on the login activity of others (e.g., by encouraging friends to use different service features). The IRFs show that customers acquired through advertising contribute the most to the firm's performance. Using these results, and assuming a discount rate $r$=0 for simplicity, we calculate the long-term multipliers (equation (4)) to be used in our value generating function as: $\gamma_{AD} = 30.51$, $\gamma_{DM} = 9.22$, $\gamma_{PR} = 16.58$, $\gamma_{WOM} = 14.03$[18]. Consistent with our expectation, we find that each of these multipliers is significantly different from the others at the 5% level, except for the difference between DM and PR[19].

[18] We do not make any assumptions on $\tau(\gamma_k)$ yet, as expressed in equation (5). Therefore, these multipliers should be interpreted as the contribution to the firm's total login activity, not to the monetary value of the firm.
[19] We tested for the differences in the cumulative impulse response function using Monte-Carlo simulations following the procedure suggested in Lutkepohl 1993 (p. 495).

**Figure 3. Impulse Response Functions**

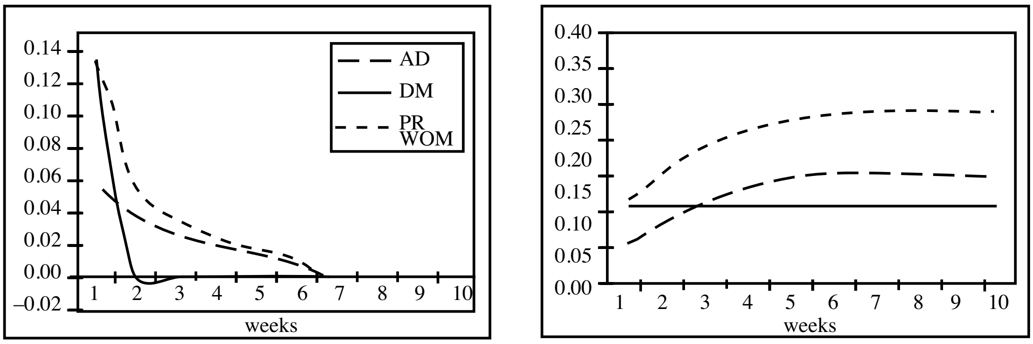
**IRF** **Accumulated IRF**

**Direct Effects**



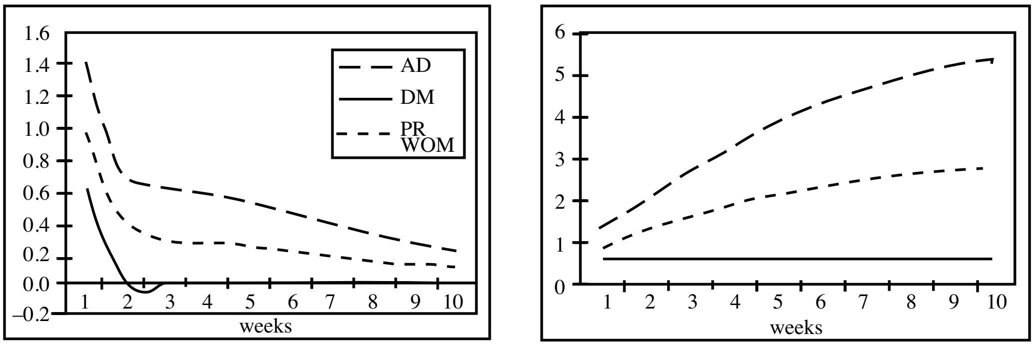

Note: Should be interpreted as the effect of one customer increase from each channel on the total login activity of the firm.

**Feedback Effects**




Note: Should be interpreted as the effect of one login activity increase on the number of customers acquired through each channel.

**Cross Effects (WOM)**




Note: Should be interpreted as the effect of one unit (customer) increase on the total number of customers acquired through word-of-mouth.

*Feedback Effects*. Here we investigate how many new customer acquisitions can be generated by an unexpected one-login increase. Indeed, increased usage of the Internet service may lead to higher customer satisfaction and reliance on the service, which can create a diffusion effect in the form of additional customer generation. The results show that increased login activity has the strongest feedback effect on public relations and word-of-mouth channels, i.e., as customers become more involved with the service, the firm enjoys higher word-of-mouth generation and also higher media coverage. By contrast, the performance feedback effect is weakest for the direct marketing and advertising channels (see Figure 4).

*Cross Effects*. We investigate only the cross effects of the different acquisition channels on the word-of-mouth channel, i.e., how effective the different acquisition channels are at generating future acquisitions through word-of-mouth. Figure 4 shows that customers acquired through advertising are better at word-of-mouth generation than those acquired in other ways. For example, each customer acquired through advertising is expected to bring around 5.4 new customers, while a customer acquired from direct marketing is expected to bring only about 0.6 customers. Surprisingly, customers acquired through word-of-mouth are less likely to generate referral customers than those acquired by advertising. These differences are managerially important and require separate research to determine their underlying causes.

*Cost-Benefit Analysis*

While the VAR model and impulse-response functions have shown that the acquisition channels generate customers of different quality (defined as contribution to CE), these channels have different acquisition costs as well. Therefore, we conduct a cost-benefit analysis in Table 6, based on published acquisition costs per channel[20]. Since this firm did not offer any referral incentives, we may set the cost of word-of-mouth acquisitions to zero. The metric "benefit per dollar" measures how many additional logins are generated by one extra dollar spent on acquiring customers in each channel. For example, while customers from advertising have the highest impact on customer equity, they are also the most expensive to acquire. In efficiency terms, advertising is found to be the least cost-effective acquisition channel.

The acquisition channels also show important efficiency differences when measured in the short run versus the long run. For example, direct marketing has a benefit per dollar around 5.2 times larger than that of advertising when measured in the short run (contemporaneous effect), but it is only 3.6 times larger when measured in the long run (total effect). Thus, managers should investigate the long-run acquisition benefits of each channel, lest they myopically favor channels with higher short-term performance but lower customer equity contribution.

[20] These costs are the averages of various industries, gathered from an independent research firm (Fiore & Collins, Successful Affiliate Marketing for Merchants, 2001).

**Table 4. Cost-Benefit Analysis**

| | *AD* | *DM* | *PR* | ***WOM*** |
|---|---|---|---|---|
| *Cost of Acquiring One More Customer ($)* | **323** | **27** | **82** | **-** |
| *Benefit: Increased Number of Logins* | | | | |
| Contemporaneous Effect | 9.07 | 3.98 | 6.21 | 5.52 |
| Lagged Effect | 21.43 | 5.24 | 10.37 | 8.51 |
| Total Effect | 30.51 | 9.22 | 16.58 | 14.03 |
| *Benefit per Dollar* | | | | |
| Contemporaneous Effect | 0.03 | 0.15 | 0.08 | |
| Lagged Effect | 0.07 | 0.19 | 0.13 | |
| Total Effect | 0.09 | 0.34 | 0.20 | |
| *Maximum WOM incentives ($)* | | | | |
| Contemporaneous Effect | 197 | 37 | 73 | - |
| Total Effect | 149 | 41 | 69 | - |

Source of cost data: Fiore & Collins, Successful Affiliate Marketing for Merchants, 2001

When word-of-mouth acquisition of customers is costless, their benefit per dollar is infinite. However, some firms implement strategies to actively boost word-of-mouth generation by referral incentives, so the question arises: What is the maximum amount a firm should be willing to pay for referrals. Table 6 provides an answer to this question by calculating the referral incentive that equates the CE contribution to that of other channels. For instance, given that the total effect of word-of-mouth on the firm's customer equity is 14.03, the firm could spend $149 per referral to obtain the same net benefit per dollar as the one exhibited by a customer acquired through advertising[21].

## A Marketing Decision Support System for Optimal Resource Allocation

The previous cost-benefit analysis rank-ordered the different acquisition channels in terms of benefit per dollar (see Table 6). This approach, while managerially insightful, has three limitations. First, it assumes that, for each channel, the acquisition cost per customer does not change with the number of customers acquired. This is equivalent to assuming that the acquisition response function is linear with no intercept. Second, it does not address the question of how much to spend on acquisition, nor does it reveal how the budget should be allocated among the different acquisition channels.

In this section, we develop a marketing decision support system (MDSS) to determine the optimal acquisition budget and its allocation across the different acquisition

[21] This reasoning assumes that customers acquired through incentivized WOM will behave in the same way as those acquired through spontaneous WOM.

channels. We have argued earlier that the objective should be to maximize customer equity, which is different from maximizing "eyeballs" or customer counts. We will show that, if the manager uses a different objective, the resulting allocation will be suboptimal. Our MDSS is similar to that of Mantrala, Sinha, and Zoltners (1992), in that we use submarket (in our case, channel) acquisition response functions to derive optimal spending and allocation across submarkets. Assuming four acquisition channels $k = \{AD, DM, PR, WOM\}$, we define a concave acquisition response function for each in the form[22],

$$c_k = \alpha_k + (S_k - \alpha_k)(1 - \exp(-\beta_k x_k)) \tag{8}$$

where $c_k$ is the number of customers acquired, $x_k$ is the amount of money spent, $S_k$ is the maximum number of customers that can be acquired (saturation level), $\beta_k$ represents the rate at which the number of customers approaches the saturation level, and $\alpha_k$ is an intercept that captures the number of customers acquired when no investment is made.

These acquisition response functions can be parameterized for each channel using decision calculus (see Blattberg and Deighton 1996 for a similar approach)[23]. The optimal resource allocation finds the best investment for each acquisition channel $x_k^*(B)$, given a fixed budget $B$. It is also possible to determine the optimal acquisition budget $B^*$ and then derive $x_k^*(B^*)$. The allocation problem can be expressed as

$$\underset{x_k(B)}{Max}\ \Pi(B) = \sum_k m_k\,(\alpha_k + (S_k - \alpha_k)(1 - \exp(-\beta_k x_k))) - B \tag{9}$$

$$s.t.\ \Sigma_k x_k \leq B,\ x_k \geq 0,\ k = \{AD, DM, PR, WOM\}$$

where $m_k$ is the contribution margin for each customer acquired from a specific acquisition channel, and $B$ is the acquisition budget. We further assume that firms exhaust their entire budget, that is $\Sigma x_k = B$.

*Incorporating Differences in the Contribution to the Firm's Profitability*

Allocations that maximize aggregate acquisitions (i.e., $\Sigma c_k$) do not necessarily maximize aggregate profits, because customers differ in their customer equity contribution. In contrast to Mantrala, Sinha, and Zoltners (1992), we incorporate the possibility of different contribution margins $m_k$ for the different submarkets (i.e., acquisition channels). If the manager's objective is to maximize customer equity, $m_k$ should represent the expected contribution of a new customer acquired through channel $k$ as explained in previous sections.

---

22 We assume that there are no cross-effects of acquisition responses among channels. This assumption could be relaxed by incorporating in each acquisition response function the effect that a certain spending in another channel will have on the number of customers acquired in that specific channel. See Rangaswamy, Sinha, and Zoltners (1990).

23 An estimation alternative to decision calculus would be using a statistical model on historical data. This may present several challenges. First, it may prove difficult to collect data on some channels such as public relations and word-of-mouth. Second, a sufficient number of data points with enough variability are required. Third, the data generation process should be able to predict future behavior. If these requirements are not met, decision calculus may be superior to statistical modeling. An example of a successful implementation of decision calculus may be found in Lodish et al. (1988).

Nevertheless, some managers have a short-term objective and want to maximize profits in the first period of the relationship. We specifically study three decision models, depending on whether the manager maximizes short-term profits or customer equity, and whether the manager takes into account the heterogeneity in the marginal value of customers or not. We show results for these three models and illustrate the effects of a short-term maximization strategy on resource allocation:

*Model 1. Same Value in the Short-Term across channels (SVST).* In this case the manager assumes that every acquired customer behaves similarly, therefore

(10)
$$\Pi(B) = m\sum_{k} \mathrm{c}_k - B$$

We explained in previous sections why this may not be a good assumption in most scenarios.

*Model 2. Different Value in the Short-Term across channels (DVST).* As shown in the empirical illustration, there is heterogeneity in customers' login activity depending on their acquisition channel. Hence, assuming that each acquisition channel brings the same "average" quality of customers may result in a suboptimal allocation. In this model we account for differences among channels, but only in the short term. We propose to use the contemporaneous impact multipliers from the IRF in the following way,

(11)
$$m_k = \tau(\phi_{vk}(0))$$

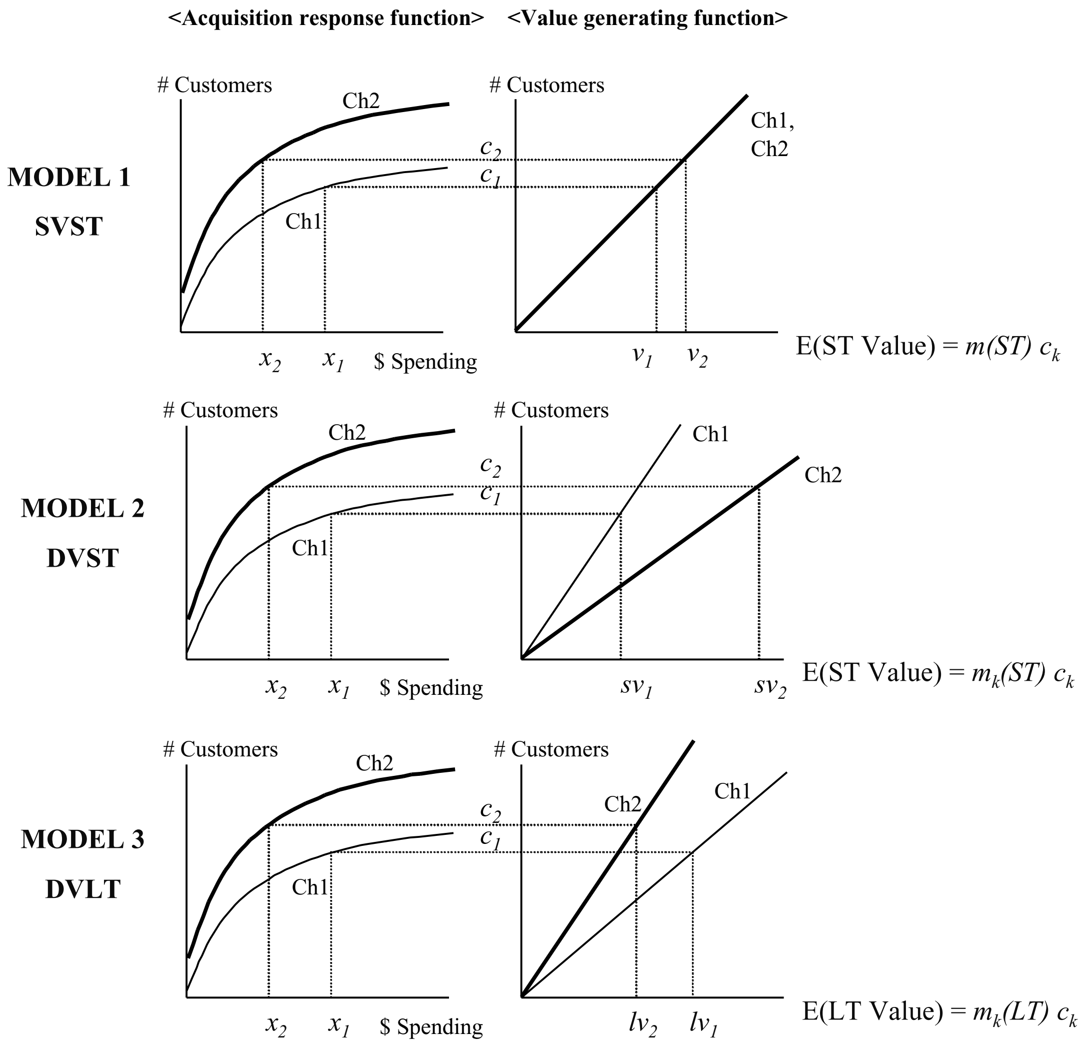
where $\phi_{vk}(0)$ is the contemporaneous impact multiplier of the direct effect of one customer acquired from acquisition channel $k$ on the firm's performance $V$. When this impact multiplier cannot be expressed as profits, a function $\tau(\gamma_k)$ should be used. For example, in our empirical illustration, we estimated the marginal contribution of an acquired customer on the firm's login activity, and we showed how login activity relates to subsequent customer revenue generation.

*Model 3. Different Value in the Long-Term across channels (DVLT).* Even though model 2 is superior to model 1 in that it accounts for the differences in the contribution margins across channels, it only incorporates differences in the contemporaneous effects. In order to obtain long-run differences, we use contributions to the firm's customer equity. For that, we will use the long-term multipliers specified as $\lambda_k$ in equation (5) such that

(12)
$$m_k \equiv \lambda_\mathrm{k} = \tau(\gamma_k)$$

Therefore, $m_k$ in this model should be interpreted as the contribution of a person acquired from acquisition channel $k$ to the firm's customer equity. That is, $m_k$ captures the contribution of one customer to both current and future profits.

**Figure 4. Acquisition Response and Value generation Functions**




where,
$x_k$ : acquisition spending for channel $k$
$c_k$ : number of customers acquired through channel $k$
$v_k$, $sv_k$, $lv_k$ : firm value generated from customers acquired through channel $k$, under models 1, 2, and 3, respectively.

Figure 6 graphically illustrates these three models. The optimal acquisition budget and its allocation logically depend not only on the acquisition response function but also on the expected monetary value that corresponds to the investment in each channel. This expected value depends, of course, on the number of customers acquired and on the assumptions made about the expected contribution from each customer as well. For example, some channels could be superior to others in the short term, but inferior in the long term. This information, captured by the value generating function[24], together with the acquisition response functions, is sufficient to derive the optimal resource allocation. Since the objective of the firm should be to maximize customer equity, we argue that model 3 (DVLT) should be superior.

*Numerical Illustration*

We provide an illustration using the results from our VAR model[25] on four different acquisition channels. For the parameterization of the value generating functions we use the contemporaneous and the long-term multipliers as reported in Table 6 and we derive customer profitability based on equations (11) and (12). Therefore, the values $\gamma_{AD} = 30.51$, $\gamma_{DM} = 9.22$, $\gamma_{PR} = 16.58$, $\gamma_{WOM} = 14.03$, $\phi_{v,AD}(0) = 9.07$, $\phi_{v,DM}(0) = 3.98$, $\phi_{v,PR}(0) = 6.21$, and $\phi_{v,WOM}(0) = 5.52$ are used. For the short-term effect of model 1, we calculate the average of the immediate multipliers, which is 6.20[26]. To calculate the acquisition response function for each channel, we use equation (8)[27].

We find the optimal acquisition budgets ($B^*$) for each of these models to be \$139,421, \$125,959 and \$387,296 respectively. *There is a substantial difference between the optimal acquisition budget for model 3 and those for models 1 and 2*. Therefore, if the manager's objective is to maximize short-term profits (using either SVST or DVST), she will underspend on acquisition. Figure 8 shows the optimal acquisition budget and allocation of the budget to each channel under different models.

[24] We assume linearity in the value generating function. This means that the expected CE contribution of a customer does not depend on the number of customers acquired from that particular channel.

[25] We are assuming that the past data generation process is valid for making predictions for the future. Under some circumstances this data would not be valid to be used for prediction purposes. For example, if the life cycle of the firm changes, or a new competitor enters into the market.

[26] Since $\tau(\gamma_k)$ is assumed common to all channels, it will not affect the optimal spending in each channel for a specific budget. Nevertheless, it will affect the optimal acquisition budget. For the purpose of this illustration we assume this function to be linear with no intercept and slope 20, which means that each login is worth 20 dollars.

[27] Equation (8) has three parameters that are obtained as follows. First, for the saturation level and the intercept we use the maximum and minimum number of registrations during these 70 weeks. We assume no intercept for advertising (AD) and direct mail (DM) since these are channels with higher possibilities of marketing intervention, whereas it is more likely to get customers without investing a penny from channels such as word-of-mouth and public relations. Second, we pick the sensitivity parameter $\beta$ so that marginal cost is equal to the average cost of the channel (as reported in Table 4) in the middle of the spending range. Our parameter values are: $\beta_{AD} = 0.0075$, $\beta_{DM} = 0.0550$, $\beta_{PR} = 0.0125$, $\beta_{WOM} = 0.0210$, $\alpha_{PR} = 103$, $\alpha_{WOM} = 67$, $s_{AD} = 550$, $s_{DM} = 953$, $s_{PR} = 1{,}104$, $s_{WOM} = 1{,}746$.

**Figure 5. Acquisition Allocation at the Optimal Budget**

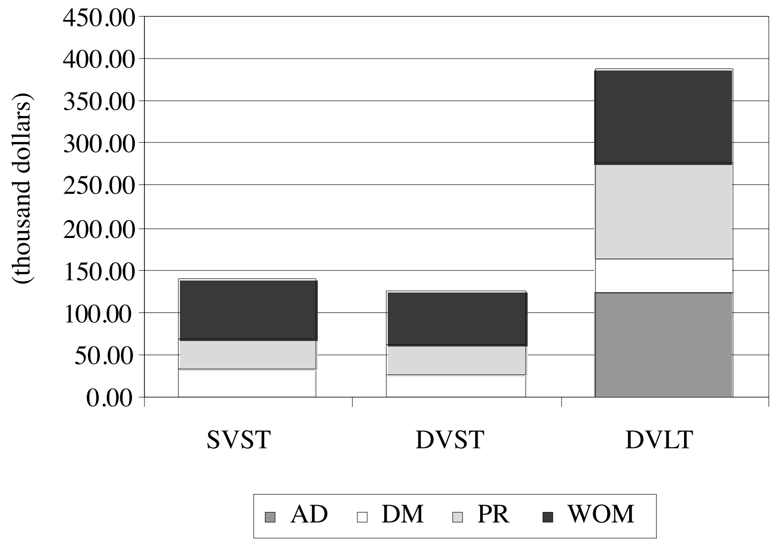
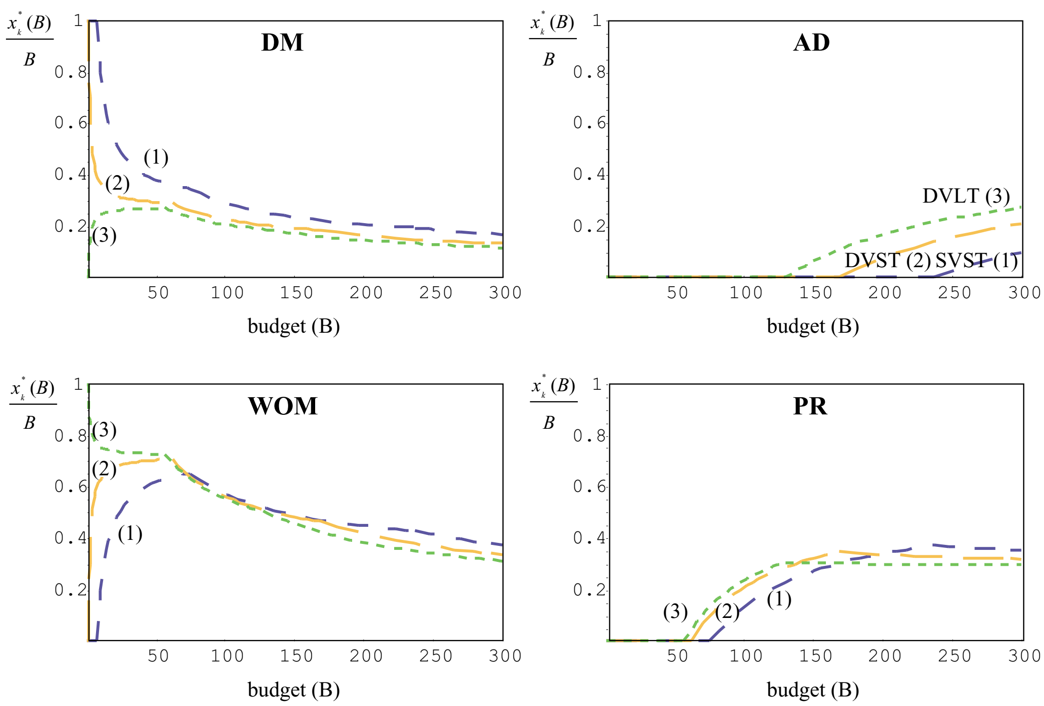



**Figure 6. Optimal resource allocation as a percentge of the acquisition budget**




Note: $x_k^*(B)$ is optinal acquisition spending in channel $k$ under the budget constraint $B$.

In many firms, the acquisition budget is set by senior management, and the marketing executive only has discretionary power over the allocation of that fixed budget. It is therefore relevant to study the optimal allocation resulting from each of our three models. Figure 10 shows the optimal allocation to each acquisition channel as a percentage of the total budget for each of the three models. *The allocations diverge substantially for small values of the acquisition budget, and tend to converge to each other for high values.* Indeed, for high values of $B$, the firm is close to the saturation level of all channels, and therefore one additional dollar spent on any channel has a small impact on profits. In contrast, for small values of $B$ any small change in the allocation across channels has a substantial impact on profitability. For example, if the firm is maximizing the contribution to customer equity (DVLT), a small budget will be spent mostly on generating word-of-mouth. However, if the firm is maximizing short-term profits (either SVST or DVST), it will spend mainly on direct marketing. Advertising is the channel that should receive the lowest allocation and firms only start to invest in advertising for sufficiently high values of $B$. When firms maximize customer equity, advertising spending begins when the budget approaches \$150,000; but when firms follow SVST, they only start to advertise when $B$ is around \$250,000.

In summary, we have developed an MDSS that can incorporate the long-run effects of each acquisition channel along with the acquisition response functions. This model allows us to determine both the optimal acquisition budget and the optimal resource allocation that maximizes customer equity. We have shown that myopically following a short-term maximization strategy will lead the manager both to underspend in acquisition and to allocate a limited budget to channels that exhibit higher short-run returns that are lower in the long-run.

## Concluding Remarks

This paper has linked a statistical model capable of measuring the long-run impact of customer acquisitions on customer equity to an MDSS that determines optimal acquisition spending and its allocation across channels. To the best of our knowledge this is the first attempt of its kind. The VAR model allowed us to measure the financial impact of an additional customer on the firm's performance ($V$). Thus, we did not explicitly measure the marketing effort (i.e., spending), but rather the result of that effort (i.e., an acquired customer) and how that acquired customer increases the customer equity of the firm. We constructed a metric called the long-term impact multiplier, which generates the intrinsic value of the "typical" customer coming from a specific acquisition channel. This metric, based on impulse response functions, not only captures the dynamic effects that a customer will exhibit in her lifetime, but also the customer's effect on other customers (e.g., generating word-of-mouth or increasing usage level). As such, our metric captures the impact of an additional customer on the customer equity of the firm.

The MDSS demonstrated the sub-optimality of acquisition-budget allocation rules that maximize the short-term profitability of the company. We showed that, when the quality of acquired customers differs across channels, the function that is maximized significantly affects the percentage of budget spent on each channel. Moreover, we showed that the smaller the budget, the larger the differences among the three allocation models.

Our measurement and optimization methods are based on a classification of customer acquisition channels that have different levels of intrusiveness and customer contact. We expect these two criteria to have an impact on customers' long-run behavior, and our empirical results confirm this expectation. Nevertheless, we do not test formal hypotheses on these relationships, which we leave as an important area for future research.

Other limitations of our work offer areas for future exploration. First, more research is needed to understand the dynamics of word-of-mouth generation. Estimating an acquisition response function could be especially difficult for word-of-mouth for two reasons: (1) for some firms it may be difficult to "incentivize" word-of-mouth and to know which is the best way to do so (e.g., offering monetary incentives to the source or to the target of word-of-mouth, or to both); (2) it may be difficult to predict customer reaction, especially when firms have never encouraged word-of-mouth before and when customers behave strategically. Second, we do not consider the resource allocation between acquisition and retention. Our MDSS could be extended to include both criteria simultaneously. We hope that this research will enhance an appreciation for the differences in customers' lifetime value and its implication for designing effective customer acquisition strategies.

Appendix A

**The relationship between login activity and willingness to pay for a fee-based service**

The empirical example offers an unusual opportunity to study the relationship between customer usage levels of a free service and their willingness to pay when the service becomes fee-based. During the 70 weeks of our observation period, customers were not charged for the web-hosting service and did not know the firm intended to change that policy later on. Two weeks after the end of our observation period, the firm announced by email that, in two months' time, users would either agree to pay subscription fees for different service levels, or face the termination of their accounts. We obtained data on which customers declined the fee for service and which ones paid fees for at least one year after the regime switch. We entertain and test the hypothesis that free-usage levels are an indication of inherent customer utility for the service and therefore predict subsequent willingness to pay.

The hypothesis is tested using a binary logit model of customer choice, using individual data on login behavior and various demographic characteristics as independent variables. Formally, we define

$$PAY = \begin{cases} \textit{1 if customer pays} \\ \textit{0 if customer abandons} \end{cases}$$

The logit model includes the following covariates:

*(1)* *LOG20:* total binary logins during the first 20 weeks of a relationship. Since we observe customers joining the firm at different points in time, we accumulate logins during their first 20 weeks of the relationship[1]. This time period is sufficient to capture a customer's level of use and interest in the service. Furthermore, it allows us to study the login behavior of a large number of customers, i.e., those who registered between week 1 and week 50 of the observation period.

*(2)* *WEEK:* week in which the customer registered. This variable allows us to test whether early adopters (customers who joined early) have a higher conversion probability than late adopters.

*(3)* *RETAILER:* 1 if retailer, 0 otherwise. Most of the firm's customers are small companies trying to advertise or sell through the Internet. Retailers are the most common business type and constituted the main target of the firm, so a priori we expect the retailer category to have higher conversion rates than others.

*(4)* *COUNTRY:* 1 if US, 0 otherwise. Although most of the firm's customers were based in the US, some were international, so this dummy variable tests for a difference in

1 Using total logins for each customer during the 70-week period would increase sample size, but make interpretation more difficult. Indeed, a sizeable percentage of registrants do not return to the site past the initial week. In the full sample we could observe, for example, a customer registering in week 70 with an average weekly login of 1, even though (s)he never returned to the site.

Appendix (continued)

conversion probability between these nationalities. Since the US was the pioneer in the commercialization of the Internet, we expect this indicator to have a positive impact.

*(5)* *EMP:* number of employees. The firm expected their service to be most suited to the needs of small firms, because of the ease of use and simplicity of its offering. Therefore, we expect larger firms to have a smaller conversion probability.

We first estimate a binary logit model on the total sample of customers who registered between weeks 1 and 50 of the observation period. Of these free-service users, only 1,030 (1.1%) chose to stay with the company after the fees were initiated. Thus, the occurrence of PAY=1 in our sample is a rare event and the logit model logically predicts that everyone will abandon the service, which results in a 98.9% correct classification rate. Nevertheless, all parameter estimates are found to be statistically significant, and our focal construct LOG20 has a positive impact on the probability of paying (see Table A.1.).

We also estimated the model with a choice-based sampling method that balances the number of paying customers and defectors (see Ben-Akiva and Lerman 1985). This technique does not yield consistent maximum-likelihood estimates of the intercept. Following Manski and Lerman (1977), we adjust the estimated intercepts for each alternative by substracting from the exogeneous maximum likelihood estimates of the intercept the constant $1n(S_g/P_g)$, where $S_g$ is the percentage of observations for alternative $g$ in the sample, and $P_g$ is the percentage of observations for alternative $g$ in the population[2].

The estimation results using the choice-based sample are reported in Table A1. The model correctly classifies 90.8% of those who terminate and 86.5% of those who agree to pay. The average predicted probability of retention for our choice-based sample is 0.475, which is very similar to the observed 0.484. Using the revised intercept, the predicted average retention probability for the population is 0.0107, which is also very close to the observed value of 0.0111.

The results support our hypothesis of a significant and positive effect of a customer's login activity on her subsequent willingness to pay. Therefore, acquisition channels with a higher level of subsequent usage (login) activity will increase the subsequent average conversion rates. The logit results are also consistent with our demographic hypotheses: customers who registered earlier, retailers, US-based firms and firms with fewer employees are more likely to be retained than others.

---

[2] Hence, for our particular estimation results, where we find an estimated intercept of –5.120, we have to revise this intercept through the following steps. We have $S_g$ = 1,030 / 2,130 = 0.4836, and $P_g$ = 1,030 / 93,119 = 0.00111. Thus we have to subtract from the estimated intercept $1n(S_g/P_g)$. Similarly, we have $S_g$ = 1,100 / 2,130 = 0.5164, and $P_g$ = 92,089 / 93,119 = 0.9889. Therefore, we should subtract from the intercept of alternative 0 the constant $1n(S_g/P_g)$ = –0.6497. The estimated new constants will be -5.12 - 3.77 = -8.90 for alternative 1, and 0 – (-0.65) = 0.65 for alternative 0. Finally, since we want to keep alternative 0 normalized to be 0, we should add the constant -0.65 to both alternatives. The resulting revised intercept will be -9.55. For an example implementing this approach for the Multinomial Logit Model see Ben-Akiva and Lerman (1985, p. 238).

Appendix (continued)

Finally, we test the relative predictive strength of customer usage levels by estimating a logit model without the demographic covariates. This model correctly classifies 90.5% of the defectors and 84.5% of the future buyers. These numbers are very close to the percentages correctly classified when other predictor variables are included in the model. In fact, the demographic variables only add 0.3 and 2.1 percentage points for correctly classified defectors and paying customers, respectively. Additionally, a model with only demographic variables as covariates classifies correctly 75.9% of the defectors and only 59.0% of the future buyers. Thus, it is login activity, and not customer demographics, that is the leading indicator of subsequent willingness to pay.

**Table A.1. Estimation Results Binary Logit Model**

| | *Total Population (N=93,119)* | | *Choice-Based Sample (N=2,130)* | |
|---|---|---|---|---|
| | *estimate* | *s.e.* | *estimate* | *s.e.* |
| *intercept* | -9.713 | (0.579)** | -5.120 | (0.730)** |
| *intercept (revised)* | | | -9.547 | |
| *log20* | 0.281 | (0.006)** | 0.324 | (0.013)** |
| *week* | -0.013 | (0.002)** | -0.020 | (0.005)** |
| *bus* | 0.763 | (0.069)** | 0.807 | (0.152)** |
| *country* | 3.440 | (0.565)** | 3.207 | (0.690)** |
| *emp* | -0.107 | (0.051)* | -0.179 | (0.089)* |
| -2 Log Likelihood | 7,194 | | 1,287 | |
| Cox & Snell R Square | 0.043 | | 0.542 | |
| Nagelkerke R Square | 0.379 | | 0.723 | |

* Significant at the 5% level; **significant at the 1% level

## REFERENCES

Anderson, Eric, and Duncan Simester (2002), "Evidence and Explanation of Long-Run Price Effects," *Paper Presented at the Marketing Science Conference*, Edmonton, Canada, 2002.

Ben-Akiva, Moshe, and Steven R. Lerman (1985), *Discrete Choice Models*. MIT Press.

Blattberg, Robert C., and John Deighton (1991), "Interactive Marketing: Exploiting the Age of Addressability," *Sloan Management Review,* 33 (1), 5-14.

—— (1996), "Manage Marketing by the Customer Equity Test," *Harvard Business Review,* 74 (4), 136-44.

Blattberg, Robert C., Gary Getz, and Jacquelyn S. Thomas (2001), *Customer Equity: Building and Managing Relationships As Valued Assets*. Boston, Massachusetts: Harvard Business School Press.

Briggs, Rex, and Nigel Hollis (1997), "Advertising on the Web: Is There Response Before Click-Through?," *Journal of Advertising Research,* 37 (2), 33-45.

Bronnenberg, Bart J., Vijay Mahajan, and Wilfried R. Vanhonacker (2000), "The Emergence of Market Structure in New Repeat Purchase Categories: The Interplay of Market Share and Retailer Distribution," *Journal of Marketing Research,* 37 (1).

Brown, Jacqueline Johnson, and Peter H. Reingen (1987), "Social Ties and Word-of-Mouth Referral Behavior," *Journal of Consumer Research,* 14 (3), 350-362.

Dekimpe, M. G., and Dominique M. Hanssens (2000), "Time-Series Models in Marketing: Past, Present, and Future," *International Journal of Research in Marketing,* 17, 183-93.

Dekimpe, Marnik G., and Dominique M. Hanssens (1995a), "Empirical Generalizations About Market Evolution and Stationarity," *Marketing Science,* 14 (3), G109-G121.

—— (1995b), "The Persistence of Marketing Effects on Sales," *Marketing Science,* 14 (1), 1-21.

—— (1999), "Sustained Spending and Persistent Response: A New Look at Long-Term Marketing Profitability," *Journal of Marketing Research,* 36 (4), 397-412.

Doan, Thomas (1992), *RATS User's Manual*. Evanston, Ill: Estima.

Doyle, Peter (2000), *Value-Based Marketing Strategies for Corporate Growth and Shareholder Value*. New York: Wiley.

Dreze, Xavier, and Fred Zufryden (1998), "Is Internet Advertising Ready for Prime Time?," *Journal of Advertising Research,* 38 (3), 7-18.

Enders, Walter (1994), *Applied Econometric Times Series*. New York: John Wiley & Sons.

Forrester (2001), "Effective Email Marketing," *hhtp://www.forrester.com*.

Friestad, Marian, and Peter Wright (1995), "Persuasion Knowledge: Lay People's and Researchers' Beliefs About the Psychology of Advertising," *Journal of Consumer Research,* 22 (1), 62-74.

—— (1994), "The Persuasion Knowledge Model: How People Cope With Persuasion Attempts," *Journal of Consumer Research,* 21 (1), 1-31.

Greyser, Stephen A., and H. Paul Root (1999), "Improving Advertising Budgeting," *MSI Working Paper,* Report No. 00-126.

Gupta, Sunil, Donald R. Lehman, and Jennifer Ames Stuart (2002), "Valuing Customers," *Working Paper at The Teradata Center for Customer Relationship Management*.

Hansotia, Behram J., and Paul Wang (1997), "Analytical Challenges in Customer Acquisition," *Journal of Direct Marketing,* 11 (2), 7-19.

Herr, Paul M., Frank R. Kardes, and John Kim (1991), "Effects of Word-of-Mouth and Product-Attribute Information on Persuasion: An Accessibility-Diagnosticity Perspective," *Journal of Consumer Research,* 17 (4), 454-62.

Jedidi, Kamel, Carl F. Mela, and Sunil Gupta (1999), "Managing Advertising and Promotion for Long-Run Profitability," *Marketing Science,* 18 (1), 1-22.

Keil, Sev K., David Reibstein, and Dick R. Wittink (2001), "The Impact of business objectives and the time horizon of performance evaluation on pricing behavior," *International Journal of Research in Marketing,* (18), 67-81.

Kwiatkowski, Denis, Peter C.B. Phillips, Peter Schmidt, and Yongcheol Shin (1992), "Testing the Null Hypothesis of Stationarity Against the Alternative of a Unit Root," *Journal of Econometrics*, (54), 159-178.`

Koop, Gary, M. Hashem Pesaran, and Simon M. Potter (1996), "Impulse Response Analysis in Nonlinear Multivariate Models," *Journal of Econometrics,* 74 (1), 119-47.

Lodish, Leonard M., Magid M. Abraham, Jeanne Livelsberger, Beth Lubetkin*, and others* (1995), "A Summary of Fifty-Five In-Market Experimental Estimates of the Long-Term Effect of TV Advertising," *Marketing Science,* 14 (3), G133-G140.

Lutkepohl, Helmut (1993), *Introduction to Multiple Time Series Analysis*. Heidelberg: Springer-Verlag.

Manski, Charles, and S. Lerman (1977), "The Estimation of Choice Probabilities from Choice-Based Samples," *Econometrica,* 45, 1977-88.

Mantrala, Murali K., Prabhakant Sinha, and Andris A. Zoltners (1992), "Impact of Resource Allocation Rules on Marketing Investment-Level Decisions and Profitability," *Journal of Marketing Research,* 29 (2), 162-75.

Mela, Carl F., Sunil Gupta, and Donald R. Lehmann (1997), "The Long-Term Impact of Promotion and Advertising on Consumer Brand Choice," *Journal of Marketing Research,* 34 (2), 248-61.

Nijs, Vincent R., Marnik G. Dekimpe, Jan-Benedict E. M. Steenkamp, and Dominque M. Hanssens (2001), "The Category-Demand Effects of Price Promotions," *Marketing Science,* 20 (1), 1-22.

Pauwels, Koen, Dominique M. Hanssens, and S. Siddarth (2002), "The Long-Term Effects of Price Promotions on Category Incidence, Brand Choice and Purchase Quantity," *Journal of Marketing Research,* 39, 421-39.

Pauwels, Koen, Jorge Silva-Risso, Shuba Srinivasan, and Dominique M. Hanssens (2003), "The Long-Term Impact of New-Product Introductions and Promotions On Financial Performance and Firm Value," *Working Paper at The Anderson School at UCLA*.

Pesaran, H. Hashem, and Youngcheol Shin (1998), "Generalized Impulse Response Analysis in Linear Multivariate Models," *Economic Letters,* 58 (1), 17-29.

Rangaswamy, Arvind, Prabhakant Sinha, and Andris Zoltners (1990), "An Integrated Model-Based Approach for Sales Force Structuring," *Marketing Science,* 9 (4), 279-98.

Reichheld, Frederick F. (1993), "Loyalty-Based Management," *Harvard Business Review,* 71 (2), 64-73.

Rust, Roland T., Valarie A. Zeithaml, and Katherine N. Lemon (2000), *Driving Customer Equity: How Customer Lifetime Value Is Reshaping Corporate Strategy*. New York: Free Press.

Sims, Christopher (1980 ), "Macroeconomics and Reality," *Econometrica,* 48, 1-49.

Srinivasan, Shuba, Frank M. Bass, and Peter Popkowski (2000), "Market Share Response and Competitive Interaction: The Impact of Temporary, Evolving and Structural Changes in Prices," *International Journal of Research in Marketing*, 17 (4), 281-305.

Thomas, Jacquelyn S. (2001), "A Methodology for Linking Customer Acquisition to Customer Retention," *Journal of Marketing Research,* 38 (2), 262-68.