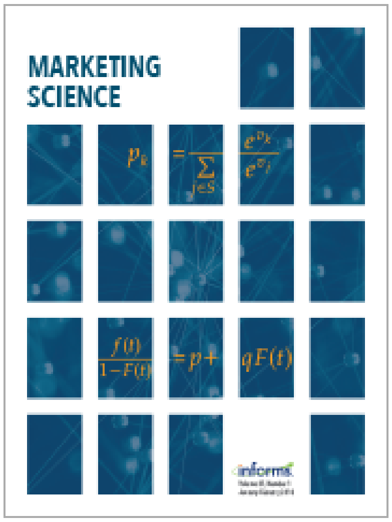


Marketing Science

Publication details, including instructions for authors and subscription information:
http://pubsonline.informs.org

Assessing the Effect of Marketing Investments in a Business Marketing Context

V. Kumar, S. Sriram, Anita Luo, Pradeep K. Chintagunta,

MARKETING SCIENCE
CELEBRATING 30 YEARS
Vol. 30, No. 5, September–October 2011, pp. 924–940
ISSN 0732-2399 | EISSN 1526-548X | 11 | 3005 | 0924

informs®
http://dx.doi.org/10.1287/mksc.1110.0661

Assessing the Effect of Marketing Investments in a Business Marketing Context

V. Kumar
J. Mack Robinson College of Business, Georgia State University, Atlanta, Georgia 30303, vk@gsu.edu

S. Sriram
Ross School of Business, University of Michigan, Ann Arbor, Michigan 48109, ssrira@umich.edu

Anita Luo
J. Mack Robinson College of Business, Georgia State University, Atlanta, Georgia 30303, aluo@gsu.edu

Pradeep K. Chintagunta
Booth School of Business, University of Chicago, Chicago, Illinois 60637,
pradeep.chintagunta@chicagobooth.edu

Recent research has empirically characterized the buyer–seller relationship as dynamically evolving from one discrete state to another. Conventional wisdom would suggest that a customer in a higher relationship state that has a higher transaction value would also have greater lifetime value to the firm. However, recent evidence suggests that higher relationship states can be ephemeral. Hence, the link between transaction value and lifetime value is not obvious. In this study, we seek to understand, within a specific empirical context, (i) the relationship between a customer's transaction value and that customer's lifetime value and (ii) the relationship between the lifetime value of a customer and the optimal level of marketing activity that needs to be directed at that customer. To this end, we develop a trivariate Tobit hidden Markov model that allows for (a) transitions among relationship states, (b) possible synergies between the various products that the supplier firm offers, (c) endogeneity in marketing activity, (d) heterogeneity in model parameters, and (e) the presence of the no-purchase option. Our results reinforce recent findings by Schweidel et al. [Schweidel, D. A., E. T. Bradlow, P. S. Fader. 2011. Portfolio dynamics for customers of a multiservice provider. *Management Sci.* **57**(3) 471–486] that higher relationship states can be short-lived. Importantly for the supplier firm, a customer in the highest relationship state in a given period does not yield the highest lifetime value to the firm. Hence, the relationship between transaction value (i.e., relationship state) and lifetime value can be nonmonotonic. At the same time, we also find a nonmonotonic relationship between the optimal expenditures that should be directed at a customer and that customer's lifetime value; i.e., the optimal level of marketing contacts is not the highest for customers with the highest lifetime value. Furthermore, we find that the optimal marketing expenditures for myopic agents are 14%–33% lower than the corresponding values for forward-looking agents. Therefore, not accounting for the long-term effects of marketing contacts would lead to suboptimal marketing budgets. Moreover, a comparison with the current marketing expenditures suggests that the current practice is closer to the myopic policy than to the forward-looking one.

Key words: business-to-business marketing; hidden Markov model; optimal resource allocation
History: Received: May 11, 2010; accepted: May 16, 2011; Eric Bradlow and then Preyas Desai served as the editor-in-chief for this article. Published online in *Articles in Advance* August 25, 2011.

1. Introduction

As firms are increasingly forced to cut costs to generate higher profits, effective allocation of scarce marketing budgets for contacting/selling to customers[1] is gaining significance. Making allocation decisions requires a good understanding of how marketing dollars affect a buyer's behavior. Resource allocation needs to be based on long-term profitability, i.e., lifetime value of a customer, rather than on one-time transactional sales or profits (Blattberg and Deighton 1996, Berger and Bechwati 2001). Key inputs into the lifetime value computation are the probability that a customer will stay with the firm (be "alive") in the future (the retention rate) and the profit stream that would accrue conditional on staying.

Recently, researchers have recognized that a dichotomous classification of customers as being "alive" or "dead" may not fully characterize the nature of their relationships with their suppliers. Rather, a customer can be in one of many relationship states at a given point in time (Oliver 1996). A nat-

[1] We use the term "customer" to refer to a business customer, a service client, or an end consumer.

ural approach to characterizing such demand is the hidden Markov model—HMM (Liechty et al. 2003, Montgomery et al. 2004, Moon et al. 2007, Netzer et al. 2008, Schweidel et al. 2011). The HMM takes a dynamic perspective and allows customers to transition from one relationship state to another. As the purchase behavior of customers varies across relationship states, the model accommodates temporal variation in the customers' frequency and quantity of purchases as they transition from one relationship state to another. A key benefit of the framework is that allowing marketing activities of firms to influence transitions from one relationship state to another explicitly allows for the long-term effects of marketing activities. This feature is critical for our empirical application in a business-to-business (B2B) context, as it has been documented that marketing dollars can help build long-term buyer–seller relationships (see, for example, Crosby et al. 1990, Sharma et al. 1999, Weitz and Bradford 1999). Because the strength of buyer–seller relationships is an enduring construct and is likely to carry over to the future (Ganesan 1994), one would expect marketing dollars to have a long-term effect on buyer behavior.

Although the notion of relationship states in hidden Markov models appears to be related to the lifetime value of customers, the literature has not carefully examined the nature of the link between the two. Specifically, if customers "live" in different relationship states associated with different levels of expenditures and frequency of purchasing, and if marketing activities influence the transitions of customers across these states, two important questions arise regarding customer lifetime value. First, if customers in higher relationship states tend to purchase more frequently and/or in greater volumes (i.e., have higher *transaction value*), are the relationship states synonymous with customer *lifetime value*? The answer to this question would depend on the difference in purchase behavior across states and the extent to which the relationship states are sticky (i.e., a high probability that a customer in a high relationship state stays in that state). Whereas the former is an empirical question, one can turn to theory for some guidance on the latter. Researchers have argued that customers who have high commitment are also likely to "seek greater relationship expansion and enhancement" (Bendapudi and Berry 1997, p. 31). If one views high purchase volumes and strong buyer–seller relationship states as evidence of high commitment, we would expect these high relationship states to be fairly sticky. Under such a scenario, one would expect a positive association between the relationship level and customer lifetime value; researchers have documented a positive relationship between purchase quantity (a proxy for relationship state) and customer lifetime value in business-to-consumer transactions (Bolton and Lemon 1999, Reinartz and Kumar 2003). However, Schweidel et al. (2011) have recently shown that the high relationship state can be very ephemeral, with customers having a high probability of transitioning directly to the "death" state. One would then expect a nonmonotonic relationship between the relationship state and lifetime value.

The second question pertains to the optimal allocation of marketing resources based on the relationship states. Specifically, *should customers in high relationship states receive more marketing investments*, or would it be optimal to spend more on the customers in lower relationship states? In addition to depending on how customer lifetime value varies as a function of relationship states, the answer to this question would also depend on how marketing contacts influence purchase behavior over time. Again, one can draw some general inferences based on the extant literature. For example, Reinartz and Kumar (2000) show that large volume buyers tend to require greater amounts of resources directed at them. If this is true, it might be optimal to focus more resources on these customers. Moreover, their purchase volumes are likely to justify greater marketing expenditures. On the other hand, if the cost of retaining these customers at the high relationship state exceeds the benefit of doing so, the firm would be better off redirecting its resources toward customers at lower relationship states.

Our objective in this paper is to answer the two questions above in a specific B2B context. We perform our empirical analysis on purchase transactions between several small and midsize firms and a large multinational supplier of technology products. Customers purchase from this manufacturer in two broad categories of technology products—hardware and software products—and potentially perceive synergies in purchasing across the two categories. We observe data on hardware and software purchases for a sample of customer firms over time, along with the associated marketing contacts to these firms. Marketing contacts by the firm have a direct effect on purchasing as well as an indirect effect by inducing transitions among relationship states, thereby having a long-term effect on demand. Furthermore, these marketing contacts are endogenously determined, and this endogeneity needs to be accounted for in the analysis. We propose an econometric framework based on a trivariate Tobit hidden Markov model to study the effects of marketing contacts on purchasing in the two categories as well as the factors influencing the levels of marketing contacts. We then use the model estimates to compute the optimal closed-loop marketing expenditures directed at a customer as a function of the probability that it exists in each of the relationship states.

The results from the empirical analysis reveal that buyers in our data exist in three levels of relationship states, with buyers in higher relationship states likely to make purchases more frequently and in greater quantity. In other words, the transaction value of customers in the highest relationship state is the largest. We find that in the absence of marketing contact investments, the lowest relationship is the most stable, with low stickiness in the higher relationship states. However, with sufficient marketing contact expenditures, it is possible to retain a customer in the mid-level relationship state. On the other hand, as in Schweidel et al. (2011), we find that the highest relationship state is very ephemeral, with customers having a high probability of transitioning back to the lowest relationship state even after significant marketing investments. Consequently, we find that the customers in the mid-level relationship state have the highest lifetime value. However, the highest marketing expenditures are optimal for customers in the highest relationship state even though they have a lower lifetime value. Such high expenditures would ensure that they do not slide back to the lowest relationship state but rather stay in the mid-level relationship state.

The contributions of this paper are primarily substantive in nature. Specifically, we underscore and clarify the relationship between transaction value and lifetime value in a situation where there is uncertainty regarding which relationship state a customer may be in the future. In our case, we find that customers in the highest relationship state—i.e., those with the highest transaction values—do not have the highest lifetime values. Furthermore, in terms of resource allocation, the highest marketing resources need not be allocated to customers with the highest lifetime values if the marginal benefit of increasing spending is higher for a different group of customers. In our case, although customers with intermediate transaction values have the highest lifetime values, it is still optimal for the firm to allocate greater resources to customers with the highest transaction values, because this would prevent these customers from sliding into the death state. In addition, we compare the optimal budget allocation with a scenario that considers only current profit maximization and quantify the amount by which myopic agents would allocate lower than optimal resources. In our application, we find the optimal budget for myopic agents would be 14%–33% lower than those for forward-looking ones. A comparison with the current marketing expenditures suggests that the current practice is closer to the myopic policy than to the forward-looking one.

Our study also makes some progress on the methodological dimension. In particular, we extend the application of the HMM to include a buyer firm's purchase quantity decisions in addition to purchase timing and category choice in a single unified model. Second, we broaden the application of the HMM to situations where buyers make purchases across multiple product categories. In doing so, we allow for possible synergies between categories. Third, we explicitly account for endogeneity of marketing contacts in our application, an issue that the previous literature has not integrated into the hidden Markov model-based analysis. Moreover, different from previous approaches to dealing with endogeneity, we use information from the firm on its marketing policies to specify our empirical model.

1.1. Related Literature

The modeling approach and contributions of this research are related to two recent papers. Montoya et al. (2010) study optimal allocation of sampling and detailing resources in the pharmaceutical industry. The key similarities of our study and that by Montoya et al. are that both studies consider the optimal allocation of marketing resources. At the same time, there are some important differences. First, unlike Montoya et al., our resource allocation situation involves purchases in multiple categories with customers perceiving synergies between them.[2] Second, we explicitly account for endogeneity of marketing contacts in our application, which the previous literature has not integrated into the hidden Markov model-based analysis.[3]

Schweidel et al. (2011) model the evolution of telecommunications service portfolios of customers. The common features shared by our study and that of Schweidel et al. are that both studies consider purchases in multiple categories. Furthermore, both studies document that high transaction value (i.e., being in the highest relationship state) need not be synonymous with high lifetime value. Nevertheless, our study extends these common attributes in three different ways. First, in our model of purchases across multiple categories, we consider possible synergies between the categories. This is in contrast to Schweidel et al., who model the utility of a portfolio as an additive function of its individual components. Second, we further investigate the link between lifetime value and transaction value to understand how optimal resource allocation would depend on these two factors. Because the allocation of marketing resources is a key managerial control variable,

[2] Moreover, we relax the assumption in Montaya et al. (2010) that only transitions to adjacent states are feasible by using an unrestricted transition matrix.

[3] Montoya et al. (2010) test for such endogeneity but cannot reject the null hypothesis of exogeneity.

understanding this link is likely to have important managerial implications. Finally, our analysis explicitly accounts for the endogeneity of marketing instruments in the analysis.

In addition, our application context differs from these two studies in that we look at a business-to-business marketing situation. Although our demand model is general enough to be applied to a variety of contexts beyond B2B situations, our substantive findings enhance our understanding of marketing expenditures in this less researched area.

2. Model

First, we estimate the demand model for hardware and software purchases that accounts for (a) zero and nonzero purchases in either category in a given period, (b) possible synergy between categories, and (c) potential endogeneity of the marketing contacts directed at each customer. In the second step, we compute the optimal closed-loop resource allocation strategy profiles conditional on the demand parameters. In this section, we discuss the formulation of the demand model and return to the resource allocation step after we discuss the estimation results from the demand model.

2.1. Overview of the Demand Model

Our supplier firm markets products in several categories (hardware and software), and customer firms make purchases across these categories. Moreover, purchases across categories may be related because of possible synergies between them or because of budget constraints. Another consequence of the presence of multiple products within each category (servers, printers, storage units, peripherals, etc., in hardware) is that it is a difficult task to model demand for each product separately. Given the diversity across products in their characteristics and prices, we instead model a customer's expenditures at the product category level (hardware/software) rather than the demand for products within each category (see also Du and Kamakura 2008 for a model of household expenditure decisions).

We propose a trivariate Tobit HMM in which the first two equations characterize customer expenditures in the two categories. The third equation characterizes the supplier firm's investments in marketing dollars directed at each customer during each period and accounts for the endogeneity of the marketing expenditure decision by the supplier firm (Villas-Boas and Winer 1999). The HMM approach models transitions between discrete latent states as a Markov process (MacDonald and Zucchini 1997). In our model, we allow marketing investments to have a direct, contemporaneous effect on customer purchases. Furthermore, marketing dollars also have an indirect effect by influencing transitions among relationship levels. Because the current buyer–seller relationship levels can influence future levels, marketing dollars have a long-term effect on expenditures.

2.2. The Hidden Markov Model of the Strength of the Buyer–Seller Relationship

The transition matrix (Q) in the hidden Markov model determines the probability of the relationship switching between the various states. Because the data on a given customer are typically "left-censored," we need to make an assumption regarding the initial state distribution (π), i.e., the probability that a customer firm is in each of the states prior to the observed data. The model also includes a vector of probabilities that relate the latent states to the observed expenditures ($\mathbf{Y}$). At any given time t, let K_{it} denote the strength of the relationship between the supplier firm and its customer firm i at time t. This relationship can be in one of NS, where k refers to the "relationship state," which could be $1, 2, \ldots, NS$ discrete states. M_{it} denotes the marketing dollars buyer firm i received at time t. The marketing dollars influence the transitions from one relationship state to another. We denote the category of interest as c, where $c = 1, 2, \ldots, C$ ($C = 2$ in our case).

2.2.1. The Transition Matrix. The HMM approach allows for transitions among the levels of the buyer–seller relationship in discrete levels or states. Correspondingly, in each period, the buyer–seller relationship can stay at the same level as in previous periods or transition to a higher or lower level. To allow transitions to all possible states, as in Schweidel et al. (2011), we use a multinomial logit model to characterize the transition matrix. Stated formally, for a model with S relationship states, we define the probability that buyer i moves from relationship state k in period $t-1$ to state k' in period t as

$$\begin{aligned} q_{it,kk'} &= \Pr(K_{it} = k' \mid K_{it-1} = k) \\ &= \frac{\exp\big((\tilde{\alpha}_{kk'} - \tilde{\alpha}_{k1}) + (\tilde{\lambda}_{k'} - \tilde{\lambda}_1) g(M_{it})\big)}{1 + \sum_{k'=2}^{S} \exp\big((\tilde{\alpha}_{ks} - \tilde{\alpha}_{k1}) + (\tilde{\lambda}_s - \tilde{\lambda}_1) g(M_{it})\big)} \\ &= \frac{\exp(\alpha_{kk'} + \lambda_{k'} g(M_{it}))}{1 + \sum_{k'=2}^{S} \exp(\alpha_{ks} + \lambda_s g(M_{it}))}, \end{aligned} \tag{1}$$

where $\alpha_{kk'}$ is the intrinsic value of the transition from state j to m compared with the transition to state 1, $g(M_{it})$ is a function of the marketing expenditures directed at firm i at time t, and $\lambda_{k'}$ denotes the marginal effect of a marketing dollar in transitioning to state k' relative to transitioning to state 1.

The intrinsic values for transitioning from one state to another (the terms $\alpha_{kk'}$ in Equation (1)) determine the transition probabilities in the absence of marketing expenditures. We conjecture that marketing dollars from the supplier firm are likely to build buyer–seller relationships by moving them to higher states or by preventing them from falling to lower states. As a result, we expect the parameters $\lambda_{k'}$ in Equation (1) to be positive. Together, the parameters $\alpha_{kk'}$ and $\lambda_{k'}$ in Equation (1) quantify the marketing dollars required to transition the buyer–seller relationship to a higher state. These parameters will also help in understanding the relative amount of marketing contact efforts required in acquiring new buyers versus retaining existing ones. Third, depending on the values of $\alpha_{kk'}$ and $\lambda_{k'}$, there can be a significant probability that the buyer–seller relationship remains at the elevated state even after marketing contacts are reduced.[4]

2.2.2. The Initial State Distribution. We empirically estimate the probability of being in each state initially through a logit model. This is similar to the approach used previously by, among others, Roy et al. (1996). Stated formally, we define the probability that a customer belongs to each of the S relationship states at the beginning of the data as

$$\begin{aligned} \Pr(K_{i1} = s) &= \frac{\exp(\beta_s)}{1 + \sum_{s'=1}^{S-1} \exp(\beta_{s'})}, \\ \Pr(K_{i1} = S) &= \frac{1}{1 + \sum_{s'=1}^{S-1} \exp(\beta_{s'})}, \end{aligned} \tag{2}$$

where β_s, $s = 1, 2, \ldots, S$, are the parameters to be estimated.

2.2.3. Relating the State of Buyer–Seller Relationship to Observed Behavior. We present the model of quantity purchased for a two-category situation. However, the model can be readily generalized to accommodate any number of categories.[5] During each period t, buyer firm i decides whether to buy and how much to spend on the two categories. The latent utility Y^*_{ickt} that buyer i in relationship state k derives from buying in category c, where $c = 1, 2$, at time t can be expressed as

$$Y^*_{ickt} = \alpha_{ick} + \beta_{ic} f(\tau_{ict}) + X_i' \gamma_c + \delta_{ic} h(M_{it}) + \varepsilon_{ict}. \tag{3}$$

In Equation (3), α_{ick} is the amount of intrinsic value that firm i, which is in relationship state k, derives from purchasing category c; τ_{ict} is the time since buyer i made a purchase in category c; $f(\cdot)$ is a function that reflects the effect of time since the previous purchase of category c, τ_{ict}, on the buying firm i's utility for the category; and β_{ic} is the corresponding response parameter. X_i contains buyer-specific characteristics, with γ_c representing the effect of these firm-specific characteristics on the utility that a firm would derive from purchasing category c; therefore, $X_i'\gamma_c$ captures the effect of observed heterogeneity on purchase quantities. M_{it} refers to the marketing dollars invested in firm i in time t by the supplier firm; $h(\cdot)$ denotes the function of marketing dollars that has a direct, contemporaneous effect on the latent utility; and δ_{ic} denotes the effect of the marketing dollars on firm i's latent utility to purchase in category c during period t. Therefore, current marketing expenditures can have both a contemporaneous (via δ_{ic}) and long-term effect (via the relationship states described earlier). The stochastic component ε_{ict} captures firm-specific idiosyncratic characteristics that influence its utility from purchasing category c at time t. Because α_{ick} varies depending on the relationship state, we allow for different sales levels corresponding to the different discrete relationship states; i.e., when the state transitions from state k to state k' following Equations (1)–(3), the customer firm's intrinsic buying utility changes from α_{ick} to $\alpha_{ick'}$. An additional feature of the model is that $f(\tau_{ict})$ accounts for possible duration dependence in the data. The model is flexible enough to accommodate any functional form to capture this duration dependence. Together, these features of our model reflect the two main characteristics of the data described at the beginning of this section.

Although we have previously noted how we capture the effects of observed heterogeneity, unobserved heterogeneity is accounted for by allowing the intrinsic category preferences, α_{ick}, where $c = 1, 2$, and the parameters β_{ic} and δ_{ic} to be customer-specific. We assume that these heterogeneity parameters are normally distributed across customers as follows:

$$\begin{gathered} \alpha_{ick} = \alpha_{ck} + \Delta\alpha_{ic}, \quad \beta_{ic} = \beta_c + \Delta\beta_{ic}, \quad \text{and} \\ \delta_{ic} = \delta_c + \Delta\delta_{ic}, \end{gathered} \tag{4}$$

where $\Delta\alpha_{ic} \sim \mathrm{N}(0, \Psi)$; $\Delta\beta_{ic} \sim \mathrm{N}(0, \sigma^2_\beta)$; $\Delta\delta_{ic} \sim \mathrm{N}(0, \sigma^2_\delta)$; and Ψ, σ^2_β, and σ^2_δ are the variance matrix across categories and the variances of the corresponding parameters. If the latent utility for buyer i to purchase in category c at time t, Y^*_{ickt}, is greater than 0, then the buyer will make a purchase. The expected quantity for such a purchase, Y_{ickt}, is Y^*_{ickt}. If the latent utility is less than 0, the buyer will not make a purchase. A buyer decides how much to buy based on the latent utility of the choice option:

$$Y_{ickt} = \begin{cases} Y^*_{ickt}, & \text{if } Y^*_{ickt} > 0, \\ 0, & \text{if } Y^*_{ickt} \le 0. \end{cases} \tag{5}$$

[4] As we discuss subsequently, our model provides a link between the state of the buyer–seller relationship and the revenues generated.

[5] We use the terms "quantity" and "expenditure" interchangeably because our empirical application uses expenditures.

2.3. Modeling the Marketing Contact Decision

As discussed earlier, the firm's decision to allocate marketing resources across various customers might not be random; the allocation decision for marketing contacts might be related to the firm-specific idiosyncratic characteristics, ε_{ict}, in (3). Our conversations with the managers of the supplier firm revealed that they set the marketing dollars for a given customer based on their marketing budget and revenues from the previous year rather than with the explicit objective of maximizing short- or long-term profits. Forcing the observed marketing contacts to conform to a model of optimal resource allocation and imposing this in the estimation (i.e., a "full information" approach; see Chintagunta et al. 2006) might lead to biased estimates. Instead, we adopt a "limited information" approach (e.g., Nair 2007) by characterizing the marketing resource allocation decision as a flexible function of the variables that the managers consider while making their decisions. This can also be viewed as a flexible approximation of the true policy as in Nair (2007). An advantage of this approach is that while we remain agnostic about the process generating the marketing contacts, we are subsequently able to compute the profit-maximizing contact level for each customer.[6] Stated formally,

$$\begin{aligned} M_{it}^{*} &= X\eta' + e_{itk} \\ &= \eta_1 + \eta_2 M_{i,\,y(t)-1} + \eta_3 R_{i,\,y(t)-1} \\ &\quad + \eta_4 M_{i,\,y(t)-1}^2 + \eta_5 R_{i,\,y(t)-1}^2 + \eta_6 M_{i,\,y(t)-1}^3 \\ &\quad + \eta_7 R_{i,\,y(t)-1}^3 + \eta_8 Z_{it} + e_{itk}. \end{aligned} \tag{6}$$

Our data are at the monthly level. Thus the $y(t)$ subscript in Equation (6) refers to the year $y(\cdot)$ that month t is a part of, and $y(t)-1$ is the previous year. $M_{i,\,y(t-1)}$ is the marketing spending on customer i during the entire previous year. Similarly, $R_{i,\,y(t-1)}$ refers to the corresponding revenues. The term Z_{it} includes exogenous variables, such as the month of the year, which can affect the marketing expenditures directed at the customer.[7] The stochastic term allows the supplier firm to adjust spending in each month based on the knowledge of the customer's purchasing. The firm decides how many marketing dollars to spend depending on the latent utility. We observe

$$M_{it} = \begin{cases} M_{it}^{*}, & \text{if } M_{it}^{*} > 0, \\ 0, & \text{if } M_{it}^{*} \le 0. \end{cases} \tag{7}$$

We assume that the stochastic component in Equation (6) follows a normal distribution. This implies that a customer firm's expenditures follow a truncated normal distribution. The buying firm can make purchases from multiple categories (in our case, two) at any given time. To account for the possibility that the purchases in the two categories may be related to each other, and for the unobserved factors affecting the marketing spending to be correlated with purchases in these two categories, we assume that the stochastic components for hardware, software, and the marketing budget, $\varepsilon_{it} = (\varepsilon_{i1t}, \varepsilon_{i2t}, \varepsilon_{i3t})$, follow a trivariate normal distribution; i.e.,

$$\varepsilon_{it} = \begin{bmatrix} \varepsilon_{i1t} \\ \varepsilon_{i2t} \\ \varepsilon_{i3t} \end{bmatrix} \sim \mathrm{N}\left(\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_1^2 & \rho_{12}\sigma_1\sigma_2 & \rho_{13}\sigma_1\sigma_3 \\ \rho_{12}\sigma_1\sigma_2 & \sigma_2^2 & \rho_{23}\sigma_2\sigma_3 \\ \rho_{13}\sigma_1\sigma_3 & \rho_{23}\sigma_2\sigma_3 & \sigma_3^2 \end{pmatrix} \right), \tag{8}$$

where σ_c, $c = \{1, 2, 3\}$ is the standard deviation of the utility of purchasing hardware ($c = 1$), software ($c = 2$), and marketing contacts ($c = 3$), and ρ_{12}, ρ_{13}, and ρ_{23} are the respective correlations between the stochastic components.

Before we discuss the estimation of the model parameters, we note the following about the nature of complementarity or substitutability across categories. Several factors can drive correlations in purchases across multiple categories (Manchanda et al. 1999). These include heterogeneity: consumers who have a high propensity to purchase in one category may also have a high propensity to purchase in other categories (Ainslie and Rossi 1998); complementarity (or substitutability) as in, for example, Chintagunta and Haldar 1998; and coincidence: consumers purchase products together because of similar purchase cycles, physical environment in the store (Swinyard 1993), and habit (Kahn and Schmittlein 1992). In our application, we account for preference heterogeneity by allowing buyers to differ from each other in terms of their intrinsic preference to purchase each category; we control for coincidence by specifying utility as a function of the elapsed time since the previous purchase in each category. Thus, we infer the complementarity/substitutability relationship between categories through the correlation between the stochastic components of the utilities of the two categories.

Although joint purchase is not a necessary condition for inferring complementarity in the use of durable goods (see Sriram et al. 2010), buyers may perceive a benefit in purchasing the two categories together, because the cost of making the purchase (in time and effort) is shared by both categories. Such purchase complementarity can be inferred from the

[6] An alternative approach is to find a customer and time-varying instrument for the marketing expenditures.

[7] Salesforce effort is likely to vary depending on the month in the quarter and by time of year, e.g., the holiday season (Misra and Nair 2011). Thus month variables can be treated as being exogenous in our contacts equation.

contemporaneous correlation in utilities of the two categories even if they are durables.[8] In our case, we can only infer complementarity (synergy) in purchases but not in consumption. Note also that a part of the substitution effect could come from the budget constraint facing the customer firm. However, we are unable to separate out these alternative explanations from the correlation given our data.

2.4. Estimation

The estimation of the model described above requires that we specify the functional forms $g(\cdot)$, $h(\cdot)$, and $f(\cdot)$ described in Equations (2) and (3); the functions $g(\cdot)$ and $h(\cdot)$ capture the effect of marketing dollars in inducing transitions between relationship states and indirectly influencing utility respectively. The function $f(\cdot)$ captures the effect of time since the previous purchase in a given category on the utility that a customer derives in purchasing in that category. While choosing a functional form for these marketing dollars, $g(\cdot)$, we wanted to get a general sense on whether it had a monotonic or a nonmonotonic effect. To this end, we estimated a version of the model with a flexible nonparametric functional form with different effects for different ranges of marketing expenditures. The estimates from this model confirmed that the effect of marketing contacts was a monotonically increasing function within the range of our data. We then tried various monotonic functions of marketing dollars, such as linear, concave (e.g., quadratic, log), and $1/(1+M_{it})$. Based on model fit, we chose the log functional form—$\ln(1+M_{it})$.[9] Given that this is a monotonically increasing function of the amount of marketing expenditures, we expect the corresponding coefficient to be positive. We use the same functional form for $h(\cdot)$ as well. For the duration dependence function, $f(\cdot)$, we once again tried several functional forms such as monotonically increasing $f(\cdot)$ (e.g., linear, $\ln(\cdot)$) as well as nonmonotonic functions such as the quadratic function and a "nonparametric" approach using dummy variables. For our data, the flexible functional form of time dummies yielded the best fit.

The first set of parameters to be estimated, $\Theta_1 = \{\alpha_{kk'}, \lambda_{k'}\}$ for k, $k' = 1, 2, \ldots, NS$, correspond to the propensity of a buyer to transition from one relationship state to another, i.e., $\lambda_{k'}$ and $\alpha_{kk'}$ in Equation (1). The second set of parameters, $\Theta_2 = \{\alpha_{ick}, \beta_{ic}, \gamma_{ic}, \eta, \Sigma\}$, where $c = 1, 2, \ldots, C$ and $k = 1, 2, \ldots, NS$, correspond to the propensity of a buyer to make purchases in each category given that she exists in a certain relationship state. We estimate all the parameters simultaneously using a maximum likelihood estimation procedure. We impose the restriction that the intrinsic preference for category in relationship state k, $\alpha_{ck+1} = \alpha_{ck} + \exp(\Delta\alpha_{ck+1})$, where $\Delta\alpha_{ck+1}$ is a parameter to be estimated; thus customers in higher relationship states have a higher propensity to make purchases in each category.

The likelihood function comprises eight regions based on what combination of variables (hardware, software, marketing contact dollars) take on nonzero values. We use the property of the conditional distribution of multivariate normal distribution to find the likelihood of the different regions (see Appendix A of the electronic companion, available as part of the online version that can be found at http://mktsci.pubs.informs.org/). Hence, we can write the likelihood of the full model as

$$L_i(y_{i1}, \ldots, y_{iT_i}) = \sum_{k_1=1}^{NS} \sum_{k_2=1}^{NS} \cdots \sum_{k_{T_i}=1}^{NS} \left\{ \pi_{ik_1} \prod_{t=2}^{T_i} q_{itk_{t-1}k_t} \prod_{t=1}^{T_i} \prod_{r=1}^{8} L_{ritk_t}^{I_{rit}} \right\}, \quad (9)$$

where π_{ik_1} is the initial distribution of a buyer being in a relationship state, T_i is the total number of observations for buyer i, $q_{itk_{t-1}k_t}$ is the transition probability from state k_{t-1} at time $t-1$ to state k_t at time t, and L_{ritk_t}, $t = 1, \ldots, T_i$ is the likelihood corresponding to the eight possible zero and nonzero outcomes in the hardware, software, and marketing contact equations, $r = 1, 2, \ldots, 8$ (see Appendix A of the electronic companion for details).[10]

3. Data

We perform our empirical analysis using a database of business-to-business purchase transactions for the mid-market segment. The database includes monthly information on the volume of purchases made by customers from different categories of high-technology products.[11] The firm maintains records for each customer at the individual decision-making unit (DMU) level.[12] The time period of the data is from July 1999 to June 2004. Because the products differ significantly in terms of dollar value, we use the total

[8] In addition, promotional incentives offered by the supplier firm such as volume discounts could also induce simultaneous purchases. However, the firm tells us that they do not offer promotions for bundled purchases.

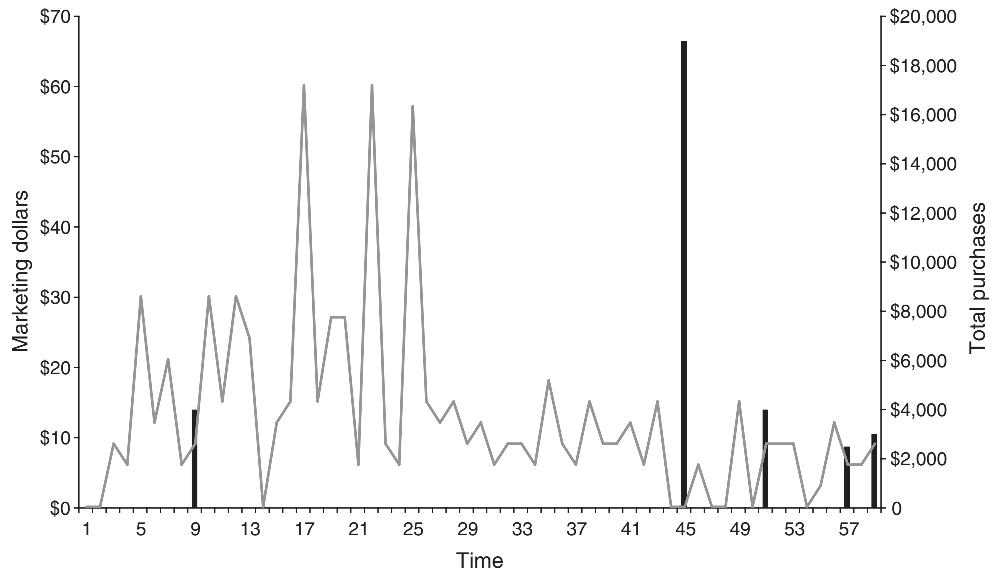
[9] This function also implies diminishing returns, which is required to obtain bounded optimal marketing expenditures on the supply side.

[10] In Appendix B of the electronic companion, we discuss identification of the model parameters based on the variation in the data. Specifically, we discuss the identification of (a) preference heterogeneity separately from complementarity, (b) the dynamics in relationship states, and (c) direct effects and indirect effects.

[11] We also aggregated the data to the quarterly level to evaluate whether the results are robust to alternative levels of aggregation.

[12] Henceforth, the term "customer" will be used to refer to an individual DMU.

Figure 1 Total Purchases and Marketing Dollars for a Representative Customer



expenditure in a given month to reflect the volume of purchases. We aggregate these into two broad categories—hardware purchases and software purchases. Furthermore, we have access to information regarding the marketing dollars that the supplier firm directed toward each buyer (i.e., DMU) in each time period. Although buyers may, at times, contact the supplier firm regarding product information, the data we have only correspond to the cost of contacts initiated by the supplier firm. Furthermore, the data do not include the contacts made by the supplier firm as a part of postpurchase support services.

To account for differences in purchase behavior across customers, we used two customer firm characteristics—the number of employees and global sales in U.S. dollars. The mean employee size is 277, and the mean global sales are about USD296 million. Importantly, the supplier firm collects the share-of-wallet information (i.e., the percentage of a buyer's information technology purchases that are made from the focal firm) for these customers on an annual basis. The average share-of-wallet across customers and over time in our data is 40%. Given that our data are for a single supplier, we use information on this variable to partially account for the effects of competition.

Recall that the objective of our analysis is to uncover dynamics in buyer–seller relationships over time as well as the role of marketing dollars in driving these dynamics. As discussed earlier, we need to observe multiple purchases for each firm in order to identify these dynamics in relationship states. Hence, for our empirical analysis, we choose a random sample of 240 firms (which is typical among the B2B firms) from a database of 506 firms that made their first purchase in the same year (i.e., 47%).[13] The strength of the data set lies in the availability of individual customer-level product purchase history as well as profile data. This allows us to use model purchase behavior at the individual firm level.

One of the unique characteristics of our data is that although the purchases at the individual customer-level are lumpy, the marketing contact efforts that induce these purchases are fairly continuous. To illustrate this point, we present the total purchases made by a representative buyer in our data set in Figure 1. The figure reveals that over the 59-month period for which we observe the buyer, there are only five occasions during which the buyer makes a purchase. On the other hand, as illustrated in Figure 1, the buyer receives marketing dollars during 50 of these 59 months. Furthermore, there is significant variation in the number of marketing dollars that this buyer receives over this period. To illustrate that this pattern can be generalized across buyers, we present the summary of marketing dollars as well as purchases made across buyers in Table 1. From Table 1, one can infer that the probability of observing a purchase by a customer in a given period is significantly smaller than the probability that the buyer will be contacted by the supplier firm.

A few other aspects of the descriptive statistics in Table 1 are worth highlighting. First, a comparison of the percentage of months during which a customer makes a purchase in either category reveals

[13] To verify that choosing a cohort did not lead to sampling bias, we compared the purchase and marketing contact profiles of two later cohorts six months apart. This did not reveal any systematic differences based on when these customers made their first purchase.

Table 1 Summary of Marketing Dollars and Purchases Across Customers

	Hardware purchase	Software purchase	Marketing dollars
% of months	2.9	1.2	44.9
Average volume (conditional) ($)	44,556	5,032	27

that hardware purchases occur more frequently than software purchases. Moreover, the average interpurchase time is approximately 19 months for hardware and 23 months for software. Additionally, conditional on purchase, the average purchase volume is higher for the hardware category than it is for software.

Because the context of our application involves customer purchases across two categories, an additional issue that we need to consider is the possible synergies in purchasing these categories together. In this regard, a relevant question is whether making a purchase in one category increases the utility from purchasing in the other category. If it does, we would expect that the average quantity of purchase in either category would be higher when the customer purchases both categories simultaneously than when he makes a purchase in only one of them. To check whether the data reveal such a pattern, we computed the average expenditures for the two categories when the firms purchase them separately and when they do so simultaneously. To account for differences in purchase behavior resulting from differing firm sizes, we classified the firms into two groups based on employment size. We present these in Table 2. Clearly, the average volumes of both hardware and software purchases seem to be significantly larger when customers make simultaneous purchases. This effect is more evident in the case of large firms. Thus, there appears to be some support for our conjecture that firms may perceive synergies when making simultaneous purchases. Nevertheless, we need to note that these descriptive statistics do not consider the (a) effect of simultaneous purchase in possibly altering interpurchase time and (b) role of marketing dollars in driving purchases in both categories simultaneously. To isolate the role of synergy from these confounding factors, we need to estimate our full model.

Table 2 Summary of Average Purchase Volumes by Firm Employee Size

	Average purchase volumes ($)			
	When buying separately		When buying together	
Employee size of firm	Hardware	Software	Hardware	Software
Small firms	34,851	2,985	34,836	4,374
Large firms	35,528	3,801	48,006	11,030

4. Results

4.1. Simulation

To evaluate the ability of the data to recover the model parameters, we carried out a simulation study. We first simulated data that mimicked those used in the analysis. Specifically, using the levels of marketing expenditures observed in the data and some prespecified parameter values, we simulated the latent utility that the customers would derive from each category (similar to Equation (3)). As in the data, we used the censored values of these latent utilities to determine the purchase quantity in each category. In the simulation, we also allowed for possible synergies between categories. Furthermore, we assumed that the marketing expenditures were set endogenously by the firm.

Next, we created 30 replications of the data and calibrated the model using each of these simulated data sets. We present the results from this simulation exercise (across all 30 replications) in Table 3. The results reveal that we are able to recover all the true parameter values precisely. We are especially able to separately identify the direct and indirect effects of marketing expenditures based on the observed variation in the data, as well as the random coefficients.

4.2. Estimation Results

We estimated the proposed trivariate Tobit hidden Markov model on the data discussed in the previous section. As in the case of latent-class models (Kamakura and Russell 1989), we need to specify the number of relationship states before estimating the model. We estimated models with two, three, and four relationship states and found that the three-state model performed the best based on Bayesian information criterion (BIC). Therefore, we restrict our subsequent discussion to the results from the three-state model.

To assess the relative performance of our model, we used three benchmark models for comparison: (a) a static version of the proposed model, i.e., a model with no relationship states; (b) a static model with marketing stock variable to capture the long-term effects of marketing expenditures;[14] and (c) a version of the proposed model but without consumer heterogeneity. The BIC for the static model with no relationship states, the static model with marketing stock variable, the HMM without heterogeneity, and the proposed model with heterogeneity were 40,798, 40,766, 40,249, and 37,286, respectively. Therefore, the three-variate Tobit HMM with heterogeneity seems to be the best among the alternatives. To assess how the proposed and the benchmark models performed in terms of out-of-sample fit, we used the

[14] We used a carryover of 0.8 while computing the stock variable.

Table 3 Random Coefficient HMM Trivariate Tobit Model Simulation

	Variable name	Estimate	Std. dev.	True value
Hardware	Intercept	2.093	0.307	2.000
	Intercept (state 2)	0.461	0.066	0.400
	Intercept (state 2)	1.353	0.104	1.400
	MKT	0.770	0.099	0.700
Software	Intercept	1.273	0.105	1.200
	Intercept (state 2)	0.121	0.187	0.100
	Intercept (state 2)	0.486	0.114	0.500
	MKT	0.413	0.126	0.400
Marketing budget	Intercept	0.293	0.103	0.300
	EMPL	0.643	0.090	0.600
Variance–covariance matrix	Hardware std. dev., exp( )	1.202	0.097	1.200
	Software std. dev., exp( )	1.020	0.122	1.000
	Marketing budget std. dev., exp( )	0.629	0.121	0.600
	Correlation of *HW* and *SW*	0.649	0.079	0.600
	Correlation of *HW* and *MKT*	0.432	0.094	0.400
	Correlation of *SW* and *MKT*	0.354	0.051	0.300
Transition matrix	Intercept (state 1 to 2)	−0.488	0.161	−0.500
	Intercept (state 1 to 3)	−0.897	0.173	−0.900
	Intercept (state 2 to 2)	−0.798	0.322	−0.700
	Intercept (state 2 to 3)	−0.989	0.245	−1.000
	Intercept (state 3 to 2)	−0.555	0.327	−0.500
	Intercept (state 3 to 3)	−0.377	0.164	−0.400
	MKT (to state 2)	0.445	0.092	0.400
	MKT (to state 3)	0.285	0.079	0.300
Heterogeneity	*HW* intercept	−0.796	0.121	−0.700
	SW intercept	−0.968	0.148	−1.000
	HW MKT	−0.483	0.204	−0.500
	SW MKT	−0.452	0.101	−0.400

Note. *EMPL*, employment size; *HW*, hardware purchase; *MKT*, marketing dollars; *SW*, software purchase.

estimated model parameters to make predictions for a holdout sample of 57 firms. The hit ratios for hardware and software purchases for the proposed model were 93% and 98%, respectively. By comparison, the corresponding hit ratios for the static model with no relationship states were 79% and 82%, respectively. The corresponding hit ratios for the static model with the marketing stock variable were 80% and 85%, respectively. Finally, the corresponding hit ratios for the HMM without heterogeneity were 82% and 89%, respectively. We also computed the relative absolute error (RAE), defined as the mean absolute deviation of a specific model relative to the mean absolute deviation of the benchmark model as proposed by Armstrong (2001). For the hardware purchase, the RAE relative to the base model—the static model with no relationship states—are 0.97, 0.90, and 0.77 for the static model with the marketing stock variables, the HMM without heterogeneity, and the HMM with heterogeneity, respectively. For the software purchase, the RAEs for the three models as mentioned above relative to the base model are 0.96, 0.89, and 0.79, respectively. The results suggest that our proposed model is able to reduce the prediction error by 21% for software and 23% for hardware compared with the base model. Overall, these results indicate that the proposed model outperforms the benchmark. The proposed HMM performs better than the model with a marketing stock variable. Although both models account for the long-term effects of marketing actions, in contrast to the model with the stock variable, the HMM allows for discrete transitions in propensity to purchase. As discussed earlier, the lumpiness of the data renders the discrete nature of the HMM more appropriate.

We present the results from the model with three relationship states in Table 4. Recall that we had imposed the restriction that customers who are in a higher relationship state have a higher propensity to purchase from the focal supplier. Therefore, the point of interest is the extent to which moving to a higher relationship state increases the propensity to purchase from the focal firm. The results in Table 5 reveal that as a customer moves from relationship state 1 to state 2, the intrinsic propensity to purchase software and hardware increases. On the other hand, when a customer moves from relationship state 2 to state 3, there is a substantial increase in his or her propensity to purchase hardware, whereas the propensity to purchase software does not appear to increase. Thus, whereas higher relationship states generally correspond to higher propensity to purchase from the focal

Table 4 Random Coefficient HMM Model Estimates

	Parameter	Estimate	Std. error	*T*-value
Hardware	Intercept	−4.429	0.108	−41.009
	Additional intercept for state 2 (exp)	−1.267	0.715	−1.772
	Additional intercept for state 3 (exp)	3.137	1.043	3.008
	Time dummy I (<7 months)	0.265	0.039	6.795
	Time dummy II (7–12 months)	0.137	0.079	1.734
	ln(1 + marketing dollar)	0.139	0.067	2.075
	ln(Gross sales)	0.029	0.011	2.636
	Employees	0.008	0.024	0.333
	Lag(Share of wallet)	0.021	0.013	1.615
	Industry 2 dummy[a]	−0.189	0.145	−1.303
	Industry 3 dummy	−0.182	0.114	−1.596
	Industry 4 dummy	−0.192	0.114	−1.684
	Industry 5 dummy	−0.087	0.097	−0.897
	Industry 6 dummy	0.085	0.128	0.664
	Industry 7 dummy	−0.215	0.119	−1.807
	Industry 8 dummy	−0.288	0.138	−2.087
	Industry 9 dummy	−0.065	0.098	−0.663
Software	Intercept	−20.056	0.307	−65.329
	Additional intercept for state 2 (exp)	3.658	0.935	3.912
	Additional intercept for state 3 (exp)	−4.298	1.025	−4.193
	Time dummy I (<7 months)	2.795	0.683	4.092
	Time dummy II (7–12 months)	0.798	0.472	1.691
	ln(1 + marketing dollar)	2.453	1.399	1.753
	ln(Gross sales)	−0.623	0.213	−2.925
	Employees	0.489	0.287	1.704
	Lag(Share of wallet)	0.314	0.176	1.784
	Industry 2 dummy	−1.899	0.452	−4.201
	Industry 3 dummy	−3.364	0.702	−4.792
	Industry 4 dummy	−2.269	0.463	−4.901
	Industry 5 dummy	−0.845	0.728	−1.161
	Industry 6 dummy	−0.099	0.935	−0.106
	Industry 7 dummy	−0.517	0.935	−0.553
	Industry 8 dummy	−6.289	1.996	−3.151
	Industry 9 dummy	−1.899	0.452	−4.201
Marketing budget	Intercept	−1.469	0.08	−18.363
	Last year's total purchase amount	0.796	0.163	4.883
	Last year's marketing expenditure	0.899	0.338	2.660
	Last year's total purchase amount^2	−0.059	0.034	−1.735
	Last year's marketing expenditure^2	−0.039	0.046	−0.848
	Last year's total purchase amount^3	0.013	0.256	0.051
	Last year's marketing expenditure^3	0.025	4.563	0.005
	Lag(Share of wallet)	0.38	0.125	3.040
Transition matrix	Parameter 1 for initial distribution	4.271	2.285	1.869
	Parameter 2 for initial distribution	1.546	0.499	3.098
	Intercept (state 1 to 2)	−3.014	0.345	−8.736
	Intercept (state 1 to 3)	−4.754	1.234	−3.853
	Intercept (state 2 to 2)	−2.676	0.654	−4.092
	Intercept (state 2 to 3)	−3.024	1.625	−1.861
	Intercept (state 3 to 2)	−3.459	1.716	−2.016
	Intercept (state 3 to 3)	−3.797	0.992	−3.828
	ln(1 + marketing dollar) (to state 2)	0.587	0.307	1.912
	ln(1 + marketing dollar) (to state 3)	0.383	0.187	2.048
Stochastic component	Hardware std. dev., exp( )	−0.046	0.053	−0.868
	Software std. dev., exp( )	2.377	0.081	29.346
	Marketing budget std. dev., exp( )	0.532	0.021	25.333
	Correlation of *HW* and *SW*	0.628	0.043	14.605
	Correlation of *HW* and *MKT*	−0.035	0.032	−1.094
	Correlation of *SW* and *MKT*	0.167	0.072	2.319

Table 4 (Cont'd.)

	Parameter	Estimate	Std. error	T-value
Heterogeneity (exp)	*HW* intercept scale	0.123	0.062	1.984
	SW intercept scale	−0.546	0.443	−1.233
	HW time dummy I (<7 months) scale	0.038	0.041	0.927
	SW time dummy I (<7 months) scale	0.0166	0.172	0.097
	HW ln(1 + marketing dollar) scale	0.06	0.039	1.538
	SW ln(1 + marketing dollar) scale	−12.3433	0.265	−46.578
Log likelihood		−18,344		

Notes. *HW*, hardware purchase; *MKT*, marketing dollars; *SW*, software purchase.
[a]The description of each industry dummy can be obtained from the authors. The industries include manufacturing, financial, retail, etc.

firm, the marginal benefit appears to vary by category. The estimates suggest that the average purchase quantities across the two categories are $12, $14,700, and $316,400 for customers belonging to states 1, 2, and 3, respectively. The corresponding gross profits (after considering the profit margins for the two categories) are $3.80, $5,874, and $44,296, respectively.[15] Therefore, the results suggest that if one considers the contemporaneous revenues, a customer in state 3 is far more valuable than a customer who is in states 1 or 2; i.e., they have *greater transaction value*. However, these results do not consider the long-term profit potential (or *lifetime value*) of customers belonging to each state.[16] Later, we discuss if this ordering is preserved when one considers the future profit potential.

The results in Table 4 reveal a positive and significant correlation between the utilities of hardware and software categories. Thus, there appear to be some positive synergies in purchasing these categories simultaneously. This reinforces the model-free empirical evidence that we found for such synergies in the data section. Such synergies could potentially create opportunities for cross selling.

The results in Table 4 reveal that both direct and indirect effects of marketing expenditures are positive and significant. Therefore, marketing expenditures seem to have a long-term effect on purchase behavior. Nevertheless, the extent to which the long-term effect persists would depend on how sticky the relationship states are. We discuss this shortly.

The results in Table 4 also reveal that there is a significant correlation between the stochastic components of the marketing budget and software purchase equations. However, the corresponding correlation with hardware purchases is not significant. Nevertheless, the results do reiterate the need to account for the endogeneity in the allocation of marketing budgets. Furthermore, the results reveal that the marketing budget allocated to a customer is indeed driven by their past purchase behavior (including purchase quantity and share of wallet) as indicated by the managers at the firm.

4.2.1. The Transition Matrix. As discussed above, when a customer moves to a higher relationship state, it is associated with a significant increase in the transaction value. However, the extent to which such a move is feasible and sustainable in the long term will depend on the values of the transition matrix as well as on the effectiveness of marketing dollars in inducing this transition. Below, we discuss these two aspects.

Transition Parameters and Transition Probabilities. The parameters of the transition matrix govern the propensity of a customer to move from one relationship state to another. We report the transition parameters for the different relationship states in Table 4. Because these parameters are difficult to interpret, we computed the implied transition probabilities in the absence of marketing dollars using Equation (3). We report these results in Table 5. The transition matrix in Table 5 reveals that in the absence of marketing dollars, there is a very small chance that a customer in a lower relationship state will move to a higher relationship state. On the other hand, a customer who is in a high relationship state has a high chance of sliding back to a lower relationship state. For example, there is a 90% chance that a customer who is in state 2 will slide back to state 1 in the absence of marketing dollars. The results are even starker for state 3, where the probability of transitioning to state 1 is marginally higher than the probability that a customer in state 1 will continue to remain in that state. This result is

Table 5 Transition Probabilities in the Absence of Marketing Dollars

	To state 1 (%)	To state 2 (%)	To state 3 (%)
From state 1	94	5	1
From state 2	90	6	4
From state 3	95	3	2

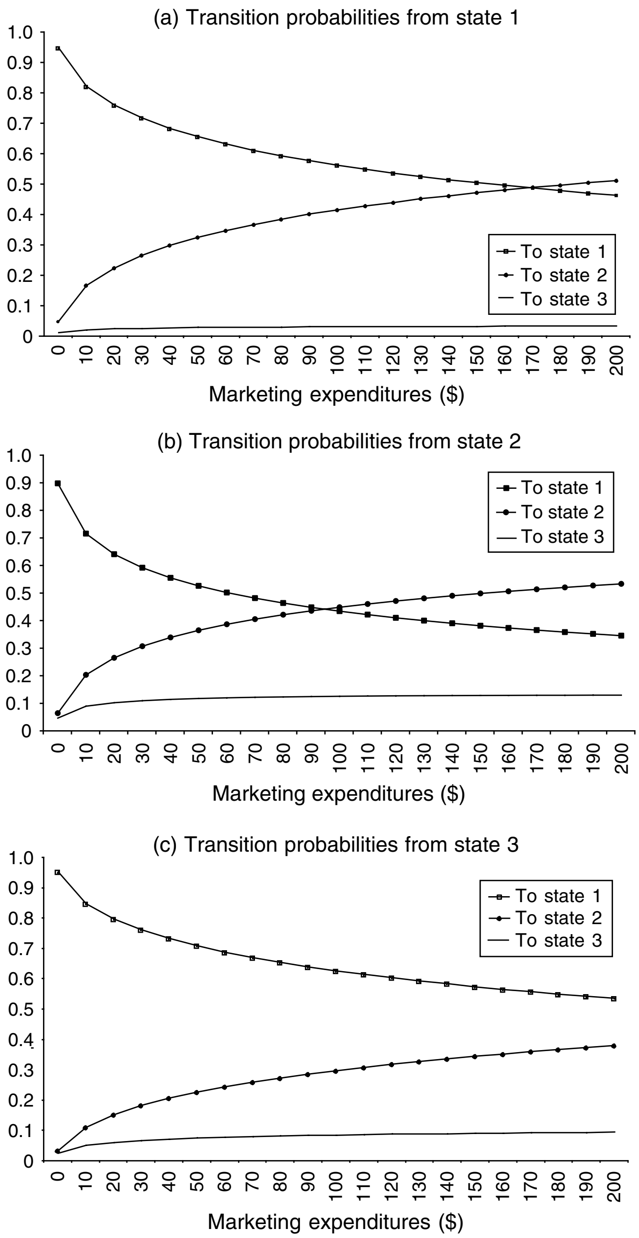
[15] We computed these values for a customer with the average employee size and annual revenue.

[16] Our discussions with the managers revealed that the average margins for hardware and software were 10% and 50%, respectively. We used these margins to convert the revenues to profits.

similar in spirit to the finding in Schweidel et al. (2011) that customers in the high relationship state can quickly transition to the lowest state. Even if such customers have the highest transaction value, they need not have the highest lifetime value. Nevertheless, taken together, the results across relationship states do imply that there is some stickiness in the states even in the absence of marketing investments.

Effect of Marketing Contacts in Inducing Transitions. To assess how these transition probabilities change with the level of marketing expenditure, we repeated them for different levels of marketing budgets. We present these results in Figure 2. The results suggest the following. First, among the three relationship

Figure 2 **Transition Probabilities as a Function of Marketing Dollars**



states, state 2 is the most responsive to marketing contacts. This is evident from the fact that the probability of staying in that state (or moving up) exceeds the probability of transitioning back to state 1 at the fastest rate for state 2. On the other hand, state 3 is the least responsive. Second, as discussed above, in the absence of marketing investments, the higher relationships (i.e., states 2 and 3) are fairly ephemeral, with high transition probabilities to the lowest state. However, when the firm commits sufficient marketing investments, state 2 can become very stable (i.e., sticky). On the other hand, state 3 is a very unstable state with high transition probabilities to state 1, even with large marketing contact investments. Once again, this is in line with Schweidel et al. (2011), who also find the highest relationship states to be transient. Together, these results suggest two things. First, given that state 2 can become very sticky with sufficient marketing investments, these investments will have significant long-term effects. Second, from the firm's point of view, it would be optimal to invest sufficient resources to keep a customer in state 2 but not so in moving them to state 3.

4.3. Optimal Marketing Resource Allocation

The stickiness in relationship states and the resulting long-term effects of marketing expenditures imply that the optimal allocation of marketing contacts is a dynamic optimization problem. Specifically, we consider closed-loop policies that would depend on the set of customer and time-varying state variables. In our context, the state variables that would determine the optimal marketing expenditures are (a) the probability that the customer exists in each of the relationship states (i.e., strength of the buyer–seller relationship), (b) the characteristics of the customer such as revenue and the number of employees, and (c) the time since the previous purchase in each category.

We consider an unconstrained (in terms of total marketing budget) optimization problem, which implies that the optimization exercise can be performed independently for each customer. Moreover, for the sake of convenience, we take a representative customer whose characteristics, such as revenue and the number of employees, are given by the average values reported in §3. Consequently, customer characteristics are not included as state variables in this discussion. Including customer characteristics as an additional state variable is a straightforward extension, especially given the unconstrained nature of optimization.

We assume that the timing of the marketing expenditure allocation decisions is as follows. At the beginning of each period, the manager observes the probability that the customer exists in each of the three relationship states, π_{1t-1}, π_{2t-1}, and π_{3t-1}. As the probabilities sum to 1, we have two independent parameters, θ_{1t} and θ_{2t}, that characterize the three states. These parameters are related to the state probabilities as follows:

$$\pi_{1t-1} = \frac{\exp(\varrho_{1t})}{1+\exp(\varrho_{1t})+\exp(\varrho_{2t})},$$
$$\pi_{2t-1} = \frac{\exp(\varrho_{2t})}{1+\exp(\varrho_{1t})+\exp(\varrho_{2t})}, \tag{10}$$
$$\pi_{3t-1} = \frac{1}{1+\exp(\varrho_{1t})+\exp(\varrho_{2t})}.$$

Recall that we model duration dependence using three time dummies: 0–6 months, 7–12 months, and >12 months. Therefore, a customer who has not purchased in the category for 6 months would transition to the 7–12 months category if he or she does not make a purchase in the current period. Likewise, a customer who makes a purchase would also reset his or her clock on the "time since the previous purchase" state to 0. Together, the transformed state membership probabilities in (10) and the time elapsed since the previous purchase in category c, π_{ct}, where $c = 1, 2$, constitute the set of state variables. Stated formally, we can define the state vector at time t as $S_t = (\varrho_{1t}, \varrho_{2t}, \pi_{1t}, \pi_{2t})$. The firm makes its marketing expenditure decisions based on the observed values of these state variables. While making this decision, the firm knows that the effect of marketing expenditures stems from two sources: (a) a direct effect on the utility that only has a contemporaneous effect and (b) an indirect effect through its impact on relationship transition matrix Q. For a given expenditure level M, we can infer the transition matrix Q based on the model parameters (using Equation (2)). Using this transition matrix and the state vector $S_t = (\varrho_{1t}, \varrho_{2t})$, the manager can determine the probability that the customer would exist in each of the three relationship states in the next period as follows:

$$\pi_{kt} \mid \pi_{1t-1}, \pi_{2t-1}, \pi_{3t-1}, Q(M) = \sum_{k'=1}^{K} \pi_{k't-1} q_{tk'k}(M), \tag{11}$$

where $q_{tk'k}(M)$ is the probability of transitioning from state k' to state k conditional on the marketing expenditure level M (Equation (2) gives the transition probability expressions). The expected state vector in the next period, $S_{t+1} = (\varrho_{1t+1}, \varrho_{2t+1}, \tau_{1t+1}, \tau_{2t+1})$, is determined as follows:

$$\varrho_{1t-1} = \ln\left(\frac{\pi_{1t}}{\pi_{3t}}\right), \tag{12a}$$

$$\varrho_{2t-1} = \ln\left(\frac{\pi_{2t}}{\pi_{3t}}\right), \tag{12b}$$

$$\tau'_{ct+1} = \begin{cases} 0, & \text{if } \{Y_c(S_t) > 0\}, \\ \tau_{ct}+1, & \text{if } \{Y_c(S_t) = 0\}, \end{cases} \tag{12c}$$

where the term $\{Yc(St) > 0\}$ represents the event that a customer makes a purchase in category c given state S_t.

Equations (3) and (5) provide the link between sales in each category and the state as well as the control variable, i.e., the marketing expenditures M. We can then determine the expected quantities of purchases in each category using the moments of the truncated bivariate normal distribution (see Rosenbaum 1961). Given these expected quantities and the corresponding margin information (we assume margins of 10% and 50% for the hardware and software categories, respectively), we can determine the expected profits as a function of the marketing expenditures, $\Xi(M)$. Given the state vector S_t and the corresponding strategy profile, the expected present discounted value of profits is

$$V(S_t \mid M) = \mathrm{E}\left[\sum_{\tau=t}^{\infty} \delta^{\tau-t} \Xi(S_\tau, M(S_\tau)) \,\Big|\, S_t\right], \tag{13}$$

where δ is the discount factor. This is the objective function that the firm needs to maximize while making the marketing expenditure decision. Note that this can also be viewed as the *lifetime value* of the customer. The value function in Equation (13) satisfies the Bellman equation:

$$V(S \mid M) = \max_{M_t>0} \{\Xi(S, M(S)) + \delta\, \mathrm{E}[V(S' \mid S, M)]\}, \tag{14}$$

where $M(S)$ corresponds to the strategy profile for the firm as a function of the state variables. The expectation on the right-hand side of Equation (14) is over the distribution of all possible realizations of the state variables conditional on the current state S and the marketing expenditures M. This uncertainty in state transitions arises because the time elapsed since the previous purchase in each category would depend on whether the customer makes a purchase in that category during the period. Because purchases in the two categories are interdependent because of cross-category synergies, we need to consider the joint probability of purchases in the two categories while computing the expected value in Equation (14). Stated formally,

$$\begin{aligned} V(S \mid M) = \max_{M_t>0} \{ &\Xi(S, M(S)) \\ &+ \delta[P_{00}(S,M)V(S'=S_{00} \mid S,M) \\ &\quad + P_{10}(S,M)V(S'=S_{10} \mid S,M)] \\ &+ P_{01}(S,M)V(S'=S_{01} \mid S,M) \\ &\quad + P_{11}(S,M)V(S'=S_{11} \mid S,M)\}, \end{aligned} \tag{15}$$

where $P_{nm}(S, M)$, $n, m = \{0, 1\}$, captures the probability of purchase $(n, m = 1)$ or nonpurchase $(n, m = 0)$ in the two categories, respectively, as a function of the state vector S and the marketing expenditures M.

Therefore, $P_{10}(S)$ corresponds to the probability of purchase in category 1 and nonpurchase in category 2. Note that we can compute the probability of these four events using Equations (3) and (5). Similarly, S_{nm}, $n, m \in \{0, 1\}$, represents the state variables corresponding to the outcome of purchase or nonpurchase in the two categories. Specifically, this captures how the time since the previous purchase would evolve (per Equation (12c)) based on corresponding purchases in the two categories. For example, a purchase in category 1 and a nonpurchase in category 2 would reset the time since the previous purchase for category 1 and advance it by one period for category 2.

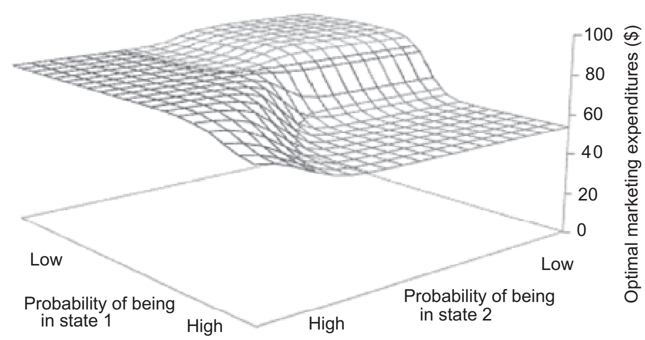
The inclusion of the duration since the previous purchase as a state variable implies that we need to compute separate value functions for different combinations of time elapsed since the previous purchase in each category. This significantly increases the computational burden of the optimization exercise. Therefore, we demonstrate how the optimization exercise can be performed for each of these duration regimes while ignoring the potential transitions. Given this simplifying assumption, the only state variables that a manager needs to consider are the probabilities that a customer would exist in any of the three states. We can then rewrite the Bellman equation (Equation (14)) as

$$V(S \mid M) = \max_{M_t > 0} \{\Xi(S, M(S)) + \delta V(S' \mid M)\}. \qquad (14')$$

We present the corresponding optimization algorithm in Appendix C of the electronic companion.

We show the results from this analysis for a customer who made a purchase in both categories in the last six months in Figure 3.[17] The results reveal the optimal level of marketing expenditures in a given period varies from \$54 (with a bootstrap standard error of \$5.8) to \$87 (with a bootstrap standard error of \$11.79). Furthermore, the results reveal that these optimal marketing expenditures vary significantly based on the probability that the customer exists in each relationship state. For example, the optimal marketing expenditures appear to be relatively high when the customer has a high probability of being in relationship states 2 and 3. In fact, the highest levels of marketing expenditures seem optimal when the probability of being in state 3 is high (i.e., low probabilities of being in states 1 and 2). On the other hand, lower levels of marketing contacts are optimal when the probability of being in state 1 is relatively high. We now comment on the result that a customer in the highest relationship state should receive the highest marketing expenditure. Recall that a customer who is in state 3 has a very high probability of transitioning to the lowest state. On the other hand, state 2 is relatively stable, with a lower probability of transitioning to the lowest relationship state. To illustrate this point, consider the probabilities of transitioning to the lowest state at different values of marketing expenditures in Figure 2, panels (b) and (c). We see that to bring down the probability of transitioning to the lowest state to 70%, the firm only needs to spend \$10 on marketing contacts if the customer is in state 2. On the other hand, if the customer is in state 3, the firm needs to spend approximately \$70 to achieve the same result. Therefore, the firm needs to direct more resources toward customers in state 3 (as opposed to those in state 2) to prevent them from sliding back to the more undesirable state 1. This explains why the optimal marketing expenditures are higher for customers in state 3.

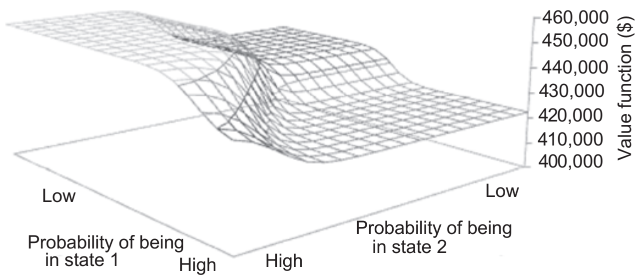
Figure 3 Optimal Marketing Contacts



An issue to consider is whether the optimal marketing expenditures would look different if we were to take a myopic rather than a forward-looking perspective. The stickiness in the relationship states would imply that a forward-looking agent who considers the long-term implications of marketing expenditures should arrive at higher optimal expenditures compared with a myopic one. At the same time, if the differences across policies are small, one could argue that the myopic approach reasonably approximates the dynamic one. To quantify the extent of this difference, we computed the optimal myopic marketing expenditures by setting the discount factor = 0. The results suggest that the optimal myopic marketing expenditures would vary between \$37 (with a bootstrap standard error of \$4.44) and \$75 (with a bootstrap standard error of \$10.41) based on the probability of existing in each of the relationship states. Thus, the optimal marketing corresponding values for forward-looking agents. Therefore, not accounting for the long-term effects of marketing contacts would lead to suboptimal marketing budgets. Moreover, a comparison with the current marketing expenditures (see Table 1) suggests that the current practice is closer to the myopic policy than to the forward-looking one.

[17] The substantive insights regarding optimal policy were similar for different combinations of duration since purchase in each category. Therefore, for the sake of brevity, we just discuss the results from one combination.

Figure 4 Value Function at the Optimal Expenditure Levels



We present the corresponding value functions (the expected net present value of all future profits) in Figure 4. Interestingly, the value function corresponding to low probabilities of being in states 1 and 2 (i.e., high probability of being in state 3) is not the highest. Rather, the value function is highest when the customer is currently in state 2 and lowest when they are in state 1. This is counter to our discussion in §4.2, indicating that the expected quantity of purchases (i.e., transaction value) for customers in state 3 was an order of magnitude greater than for those in state 2. This seemingly counterintuitive result (value function in state 2 is higher than that in state 3) can be rationalized by noting that the probabilities of being in each state correspond to those in which the customer exists at the end of the previous period. Therefore, current period payoffs would depend on how the customers are expected to transition from this state. As discussed above, while customers in state 3 have a very high probability of transitioning to the lowest relationship state, state 2 is relatively stable. Hence, when we consider the long-term profit potential, customers in state 2 are more valuable. Together, these results have two main implications. First, they suggest that customer lifetime value metrics based on past metrics such as recency, frequency, and monetary value (RFM) scoring might not be appropriate.[18] Rather, one needs a forward-looking metric that considers the future profit potential of customers (Reinartz and Kumar 2000, Gupta and Lehmann 2003). Second, customers for whom the optimal marketing expenditures are high (i.e., those with high probability of being in state 3) do not have the highest long-term profit potential (value function). This suggests that it might not always be optimal to invest in marketing resources on your most profitable customers (Bodapati 2008). Rather, we need a more careful analysis by taking into consideration the transition probabilities and the relative benefit that would accrue from moving to higher relationship states.

[18] A customer in relationship state 3 might score high on RFM-based metric but does not fare as well when one considers the future profit potential.

5. Conclusion

Although conventional wisdom would suggest that a customer in a higher relationship state that has a higher transaction value would also have greater lifetime value to the firm, recent evidence has shown that higher relationship states can be ephemeral. Hence, the link between transaction value and lifetime value is not obvious. In this study, we seek to understand, within a specific empirical context, (a) the relationship between a customer's transaction value and that customer's lifetime value and (b) the relationship between the lifetime value of a customer and the optimal level of marketing activity that needs to be directed at that customer. Our results reinforce the findings by Schweidel et al. (2011) that higher relationship states can be short-lived. Importantly, the highest relationship state does not yield the highest lifetime value to the firm. Hence, the relationship between transaction value (i.e., relationship state) and lifetime value can be nonmonotonic. Additionally, we find a nonmonotonic relationship between the optimal expenditures that should be directed at a customer and that customer's lifetime value; i.e., the optimal level of marketing contacts need not be the highest for customers with the highest lifetime value. We believe that these results can have important implications for our understanding of buyer–seller relationships in general and the optimal allocation of marketing resources in particular. Moreover, the context of our analysis, wherein the supplier firm offers multiple related products, allows us to understand perceived synergies in customer purchases across these categories. Firms can use these results to promote cross selling.

We believe that this paper carries important substantive implications for managers in B2B markets, and it also offers avenues for future research. One can potentially estimate the parameters at the individual customer–firm level in order to say something about that firm's responsiveness and to enable better targeting of effort. A key limitation of our application is the lack of customer purchase data from competing firms. Future research can extend the model presented in this paper to formally address the issue of missing share-of-wallet information along the lines of Chen and Steckel (2005). Such a model could formally take into account how competitive actions affect customer lifetime value and the optimal level of marketing expenditures. Additionally, extending our analysis to different types of marketing contacts would broaden the resource allocation task. It would also be a worthwhile exercise to implement the optimal closed-loop policy and compare the long-term profits with those under the current heuristic. In Appendix D of the electronic companion, we discuss the steps in implementing this policy. Such an application would have a worthwhile impact on the optimal allocation decisions by practicing managers.

In summary, we have taken the first steps in studying the nature of the relationship between customer firms and the marketing company in terms of the implications for marketing contacts and lifetime value. The Tobit hidden Markov model, coupled with a formal model of dynamic marketing optimization, gives us interesting insights into the nature of industrial buying.

6. Electronic Companion

An electronic companion to this paper is available as part of the online version that can be found at http://mktsci.pubs.informs.org/.

Acknowledgments

All authors contributed equally to this effort. The authors thank the Institute for the Study of Business Markets (ISBM) for financial support and the participants at the ISBM Conference 2008 for their input. They are also grateful to the editor, associate editor, and reviewers for their comments and suggestions. The usual disclaimer applies. The authors thank Renu for copyediting the manuscript.

References

Ainslie, A., P. E. Rossi. 1998. Similarities in choice behavior across product categories. *Marketing Sci.* **17**(2) 91–106.

Armstrong, J. S. 2001. Evaluating forecasting methods. J. S. Armstrong, ed. *Principles of Forecasting*. Kluwer Academic Publishers, Norwell, MA, 443–472.

Bendapudi, N., L. L. Berry. 1997. Customers' motivations for maintaining relationships with service providers. *J. Retailing* **73**(1) 15–37.

Berger, P. D., N. N. Bechwati. 2001. The allocation of promotion budget to maximize customer equity. *Internat. J. Management Sci.* **29**(1) 49–61.

Blattberg, R. C., J. Deighton. 1996. Manage marketing by the customer equity test. *Harvard Bus. Rev.* **74**(4) 136–144.

Bodapati, A. V. 2008. Recommendation systems with purchase data. *J. Marketing Res.* **45**(1) 77–93.

Bolton, R. N., K. N. Lemon. 1999. A dynamic model of customers' usage of services: Usage as an antecedent and consequence of satisfaction. *J. Marketing Res.* **36**(2) 171–186.

Chen, Y., J. Steckel. 2005. Modeling credit card "share of wallet": Solving the incomplete information problem. Working paper, Northwestern University, Evanston, IL.

Chintagunta, P. K., S. Haldar. 1998. Investigating purchase timing behavior in two related product categories. *J. Marketing Res.* **35**(1) 43–53.

Chintagunta, P. K., T. Erdem, P. E. Rossi, M. Wedel. 2006. Structural modeling in marketing: Review and assessment. *Marketing Sci.* **25**(6) 604–616.

Crosby, L. A., K. R. Evans, D. Cowles. 1990. Relationship quality in services selling: An interpersonal influence perspective. *J. Marketing* **54**(3) 68–81.

Du, R. Y., W. A. Kamakura. 2008. Where did all that money go? Understanding how customers allocate their consumption budget. *J. Marketing* **72**(6) 109–131.

Ganesan, S. 1994. Determinants of long-term orientation in buyer-seller relationships. *J. Marketing* **58**(2) 1–19.

Gupta, S., D. R. Lehmann. 2003. Customers as assets. *J. Interactive Marketing* **17**(1) 9–24.

Kahn, B. E., D. C. Schmittlein. 1992. The relationship between purchases made on promotion and shopping trip behavior. *J. Retailing* **68**(3) 294–315.

Kamakura, W. A., G. J. Russell. 1989. A probabilistic choice model for market segmentation and elasticity structure. *J. Marketing Res.* **26**(4) 379–390.

Liechty, J., R. Pieters, M. Wedel. 2003. Global and local covert visual attention: Evidence from a Bayesian hidden Markov model. *Psychometrika* **68**(4) 519–541.

MacDonald, I. L., W. Zucchini. 1997. *Hidden Markov and Other Models for Discrete-Valued Time Series*. Chapman & Hall, New York.

Manchanda, P., A. Ansari, S. Gupta. 1999. The "shopping basket": A model for multicategory purchase incidence decisions. *Marketing Sci.* **18**(2) 95–114.

Misra, S., H. S. Nair. 2011. A structural model of sales-force compensation dynamics: Estimation and field implementation. *Quant. Marketing Econom.*, ePub ahead of print March 12.

Montgomery, A. L., S. Li, K. Srinivasan, J. C. Liechty. 2004. Modeling online browsing and path analysis using clickstream data. *Marketing Sci.* **23**(4) 579–595.

Montoya, R., O. Netzer, K. Jedidi. 2010. Dynamic allocation of pharmaceutical detailing and sampling for long-term profitability. *Marketing Sci.* **29**(5) 909–924.

Moon, S., W. A. Kamakura, J. Ledolter. 2007. Estimating promotion response when competitive promotions are unobservable. *J. Marketing Res.* **44**(3) 503–515.

Nair, H. 2007. Intertemporal price discrimination with forward-looking consumers: Application to the US market for console video-games. *Quant. Marketing Econom.* **5**(3) 239–292.

Netzer, O., J. M. Lattin, V. Srinivasan. 2008. A hidden Markov model of customer relationship dynamics. *Marketing Sci.* **27**(2) 185–204.

Oliver, R. L. 1996. *Satisfaction: A Behavioral Perspective on the Consumer.* McGraw-Hill, New York.

Reinartz, W. J., V. Kumar. 2000. On the profitability of long-life customers in a noncontractual setting: An empirical investigation of implications for marketing. *J. Marketing* **64**(4) 17–35.

Reinartz, W. J., V. Kumar. 2003. The impact of customer relationship characteristics on profitable lifetime duration. *J. Marketing* **67**(1) 77–99.

Rosenbaum, S. 1961. Moments of a truncated bivariate normal distribution. *J. Royal Statist. Soc. Ser. B* **23**(2) 405–408.

Roy, R., P. K. Chintagunta, S. Haldar. 1996. A framework for investigating habits, "the hand of the past," and heterogeneity in dynamic brand choice. *Marketing Sci.* **15**(3) 280–299.

Schweidel, D. A., E. T. Bradlow, P. S. Fader. 2011. Portfolio dynamics for customers of a multiservice provider. *Management Sci.* **57**(3) 471–486.

Sharma, A., N. Tzokas, M. Saren, P. Kyziridis. 1999. Antecedents and consequences of relationship marketing: Insights from business service salespeople. *Indust. Marketing Management* **28**(6) 601–611.

Sriram, S., P. K. Chintagunta, M. K. Agarwal. 2010. Investigating consumer purchase behavior in related technology product categories. *Marketing Sci.* **29**(2) 291–314.

Swinyard, W. R. 1993. The effects of mood, involvement, and quality of store experience on shopping intentions. *J. Consumer Res.* **20**(2) 271–280.

Villas-Boas, J. M., R. S. Winer. 1999. Endogeneity in brand choice models. *Management Sci.* **45**(10) 1324–1338.

Weitz, B. A., K. D. Bradford. 1999. Personal selling and sales management: A relationship marketing perspective. *J. Acad. Marketing Sci.* **27**(2) 241–254.