Managing marketing communications with multichannel customers

Jacquelyn S. Thomas
Associate Professor of Integrated Marketing Communications McCormikc Tribune Centre Northwestern University
jakki@northwestern.edu

Ursula Y. Sullivan
Assistant Professor of Marketing Department of Business Administration University of /1/inois Urbana
uyas@uiuc.edu

The authOIS thank the specialissue consulting edita-s, William Boulding and Richard Staelin, for their helpful comments on this artide. They are also grateful fa- the assistance of Len Berry and the Center fa- Retailing Studies at Texas A&M University.

ABSTRACT

This article presents a marketing communications process that uses customer relationship management ideas for multichannel retailers. The authors describe and then demonstrate the process with enterprise-level data from a major U.S. retailer with multiple channels. On the basis of the results, the authors develop an initial marketing communications strategy for the retailer

Over the past decade, customer relationship management (CRM) has proved critical in helping firms make more money by enabling them to identify the best customers and then satisfy their needs so that they remain loyal to the firm More recently, CRM has grown increasingly complex with the proliferation of retailers expanding their channels of distribution (Cleruy 2000). This has led to the need for enterprise-level data, which are the aggregation of data gathered from all firm interactions with their customers across all channels. Given this data-rich, dynamic environment, how can a firm identify the customers who will migrate among its multiple channels and predict their migration patterns? More important, how does the firm communicate with these customers to influence their channel choices and, ultimately, their value? This research focuses on answering these questions.

Thus, this a1licle has two general objectives. First, we illuminate a process by which multichannel retailers can leverage enterprise-level data to understand and predict their customers' channel choices over time. Second, we demonstrate how the infmmation gained from this process can be used to develop strategies for tru·geting and communicating with customers in a multichannel environment.

The benefits achieved from the application of this process include increased efficiency in ma.I·keting expenditures and enhanced customer value. In the next section, we outline the general process for managing ma.I·keting communications (MARCOM) with multichannel retail customers. Subsequently, we demonstrate the application of the process using an enterprise database of a multichannel retailer. We conclude by noting the limitations of the study and ideas for further reseru·ch in this area.

THE MULTICHANNEL MARCOM PROCESS

The process of managing MARCOM in a multichannel environment begins with the identification of relevant factors that differentiate among customers who use different channels. It continues with the development of a communication strategy for existing customers, and it ends with the prediction of the right communications strategy for prospects and new customers.

Step 1: Estimate a Segment-Level Channel Choice Model

The ctitical aspect of this step of the process is not choosing the model (e.g., multinomial logit

or probit) as much as it is specifying the model. It is important that the variables in the model are factors that (1) drive channel choice, (2) help classify prospects and customers into segments, and (3) measure the efficiency of the MARCOM expenditures and activities. Channel choice research has identified customers' price expectations (e.g., Brynjolfsson and Smith 2000), the product group to be purchased (Young 2001), and convenience (Forster 2004) as factors that may lead to a specific choice among channels or stores (Fox, Montgomery, and Lodish 2004). Balasubramanian, Raghunathan, and Mahajan (2005) assert that the goals (e.g., economic, self-affirmation, socialization) a consumer tries to achieve during his or her shopping experience affect channel choices.

Although descriptors of the actual purchase are relevant, it is also important to include independent measures that can be known before a purchase so that they can be used to classify prospects into segments. The distance that a customer lives from a store is an example. Other examples of factors that may drive channel choice include switching costs and risk aversion (Dholakia, Zhao, and Dholakia 2005). If competitive data are available, both of these factors can be assessed before the customer's first purchase with the focal firm.

Another factor to consider in the channel choice decision is the extent to which prior channel choices influence current channel choices. In general, marketers have shown that prior experiences affect current choices (e.g., Boulding, Kalra, and Staelin 1999; Thomas, Blattberg, and Fox 2004). In a channels context, Shankar, Rangaswamy, and Pusateri (2001) show that a prior positive experience with a brand in a physical store can decrease price sensitivity online. Furthermore, Dholakia, Zhao, and Dholakia (2005) assert that prior channel choices affect subsequent channel choices when customers make repeat purchases. Thus, knowledge of prior purchase channels may help explain future channel choices and develop a MARCOM strategy.

In terms of MARCOM, variables that measure the number, nature, or dollar amount of communication expenditures are also important to include in the channel choice model. For example, a study of migration between the catalog and the Internet channels finds that the number of marketing communications largely predicts the buying behavior of an Internet-oriented segment (Ansari, Mela, and Neslin 2005). Other research asserts that individual- or segment-level media expenditures are essential in the development of a MARCOM strategy (Tellis 2003, p. 45).

Although thoughtful identification of the critical factors that drive channel choice is important, it is also vital to recognize that, similar to attitudes (see Eagly and Chaiken 1993), the consumer's channel choice probabilities may change over time. This change over time is referred to as nonstationarity. In empirical applications, the presence or absence of nonstationarity in choice probabilities is determined by a statistical test (Anderson and Goodman 1957; Montgomery 1969; Styan and Smith 1964). This phenomenon can be captured simply by including a variable in the choice model that indicates time or purchase occasion number.

Step 2: Assign Each Existing Customer to a Segment and Profile the Segments

After the factors associated with channel choice have been examined and the choice model has been estimated, a statistical equation can be used to assign customers to a specific segment. A profile of each segment should then be created to describe the demographics and historical behavior of the segments, including differences and similarities. Doing so not only results in a greater understanding of the kinds of customers who frequent the retailer but also helps firms target and assign new customers to current segments (as can be observed in Step 5).

Step 3: Predict the Probability of Channel Choice over Time

The purpose of this step is to anticipate the customer's channel choices in the future, thereby becoming more efficient with MARCOM activities in a multichannel environment. Various statistical models can be used for prediction. In this process, we leverage the results from Step 1 to develop a

Markov chain. A Markov chain details the probability of a customer sequentially choosing to buy from different channels over time. A firm can then determine which channels a customer is most likely to buy from in the future. Using the knowledge of what drives channel choice (i.e., the results from Step 1), a firm can then attempt to leverage MARCOM to encourage channel choices that enhance customer value.

Step 4: Develop a Segment-Specific Communications Strategy

With this step, a firm can now begin to develop a MARCOM strategy for its existing customers. To develop the strategy, the firm should consider (1) the customer types that generate the most value to the firm (e.g., catalog plus Internet customers versus bricks-and-mortar store plus Internet customers), (2) a customer's intrinsic channel choice preferences and tendencies given the current MARCOM tactics, (3) the degree to which customers respond to MARCOM and the nature of that response, and (4) the costs associated with different MARCOM activities.

Step 5: Classify First-Time Customers into Existing Segments

The objective of Step 5 is to use early information (e.g., demographics, channel of first purchase, revenues from first purchase) from prospects and first-time customers to classify them into the existing segments they most closely resemble. There are various classification and segmentation methods (e.g., discriminant analysis, chi-square automatic interaction detection [CHAID], classification and regression tree [CART], cluster analysis) that can be used in this step. Given the classification, the firm can then tailor communications that will influence purchase behavior similar to other customers in the same segment. The more elaborate the segment profiles (from Step 2) and the more detailed the data on prospects and new customers, the easier this step becomes.

Step 6: Update Segment Affiliation

Finally, as current and new customer interaction data become available, the data can then be applied to repeat the prior steps and update the segment assignments and their profiles. In particular, the Markov chain may help update the segment memberships. Note that the six steps we outline for the enhancement of MARCOM efficiency are closely aligned with the four critical actions (i.e., database creation, market segmentation, forecasting customer purchase behavior, and resource allocation) that Berger and colleagues (2002) assert are necessary for the assessment of how marketing actions affect customer value.

APPLICATION OF THE MANCOM PROCESS

Customer Database

To demonstrate the application of the multichannel communications process, we use an enterprise customer database from a major U.S. retailer. This database includes sales from the retailer's three channels: physical retail stores, catalogs, and the Internet. The data were captured using the retailer's proprietary system, which first issues a unique number to a customer and then tracks that customer each time he or she purchases an item from any of the retailer's three channels. Of the more than 4100 customers tracked for this analysis, bricks-and-mortar store-only customers constituted approximately 63% of the total, catalog-only customers constituted 11.9%, and Internet-only customers constituted 12.4%. Dual-channel customers constituted 11.9% of the orders, and three-channel users constituted approximately 1%. Descriptive statistics of customers' purchase behaviors and highlights of the differences between customers appear in Table 1. We determined all of the relative comparisons we note in the bottom portion of Table 1 using a multivariate analysis of variance and planned contrasts for which the critical significance level was at least .05. Note that information from Table 1 is helpful when trying to derive inferences about the relationship between channel usage and customer value. Thus, we also use this information in future steps.

The data cover one year's worth of purchases from only first-time buyers.[1] Although this

[1] We limit our analysis to customers who first began a relation- ship with the firm during the observation horizon because customers who began before this time may have histories preceding the observation window that affect their current behavior and this analysis.

one-year time horizon is not ideal for lifetime value assessments, which are common in CRM, this application shows that firms with even a limited enterprise database can begin to assess channel migration and develop communication strategies. As more data become available, a firm can then repeat the process and assess longer-term customer profitability.

Although data integration across channels is a significant challenge for many firms, a failure to do so could distort the firm's view of its customers (Aberdeen Group 2004). In these data, we find that if a channel were to function in a silo, it would fail to capture between 50% and 65% of a multichannel customer's total annual revenues. The magnitude of this distortion will vary by firm, but in these data, we find that it is similar for the store and catalog channels and is the greatest for the Internet channel. Thus, the necessary characteristics of the data are that (1) they are a longitudinal record, tracking all of a customer's purchase occasions, and (2) they contain elements of the purchase environments (e.g., promotions, pricing) across all three channels. For each purchase occasion, this firm's database tracks the channel from which the customer purchases, the products purchased and their categories, and the prices paid. A notable feature of this retailer is that all products are offered in all three channels, and pricing and promotions are uniform across the channels.

Step 1: Estimate a Segment-Level Channel Choice Model

We use a multinomial logit model to estimate channel choice at a given time. A list of the independent variables and their operationalizations appear in Table 2. There are a few things to note with respect to our variable selection. First, we include a quadratic term for MARCOM dollars to explore nonlinearities in the relationship between communication expenditures and channel choice. We based this decision on previous research that has identified a decreasing-returns relationship between MARCOM expenditures and customer acquisition, retention, and long-term profitability of customers (Reinartz, Thomas, and Kumar 2005). Second, we include only the most recent prior purchase as a variable in the model. We based this decision on the statistical tests that Styan and Smith (1964) outline, from which we determined that the channel choice at time (t – 1)affects the channel choice at time t.[2] Third, we include the purchase occasion variable. Using a test that Anderson and Goodman (1957) outline, we conclude that channel choice probabilities change over time (i.e., they are nonstationary).[3]

To capture unobserved heterogeneity in the choice probabilities, we estimated the model using a latent class segmentation approach as Kamakura and Russell (1989) describe. The results from estimating the Logit model appear in Table 3. Note that the latent class procedure yielded two distinct segments (Akaike information criterion for Segments 1, 2, and 3 were 1.115, .9326, and .9330, respectively). For comparing the relative impact of the factors related to channel choice, Table 3 includes the elasticities of the statistically significant variables. These elasticities aid us in Step 4 as we develop the MARCOM strategy.

Step 2: Assign Each Existing Customer to a Segment and Profile the Segments

Consistent with the latent class estimation in Step 1, we assign customers to one of the two segments on the basis of their prior choice histories. As Kamakura and Russell (1989, Eqs. 7 and 8) describe, we first assume that a customer belongs to a specific segment. Given that assumption, we compute the likelihood of each customer's channel choice history. The equation to compute this is expressed here as follows:

$$(1) \qquad L(H_k|i) = \prod_t P_j(u_{ji}, \beta_{ji}, X_{kt}),$$

2 For a zero- versus first-order test, we need a customer to have at least two purchase occasions. However, because we do not assume stationarity at this point, for those who had a longer history, we continued to test the zero-order hypothesis at different points in their life cycles. Thus, we computed chi-square statistics at several repeat purchase occasions. The chi-square values for the first, second, third, fourth, and fifth repeat purchases were 137.33, 32 92, 26.17, 28.45, and 17 19, respectively. Assuming that *p* = .995, the critical chi-square value with four degrees of freedom is 14.86; therefore, we reject the null hypothesis that states that the process is a zero-order model. We conducted a second-order test, but we could not reject the null hypothesis that it was a first-order mode

3 Given our data, we let T = 6; we applied a test of stationarity to our data and found a chi-square value of 252.85. If *p* = .995, the critical chi-square value at 30 degrees of freedom was 50.67. From this, we conclude that we can reject the null hypothesis and that in this application, the channel choice process is nonstationary

Table 1: Data Description and Comparison

Channel Usage	Sample Size	Percentage of Population	Means (Standard Errors): Number of Unique Purchase Occasions	Number of Categories Purchased in	Total Dollars Spent over Relationship ($)	Number of Items Purchased
Brick only (B)	2620	62.95	3.66 (2.44)	3.31 (2.19)	571.21 (2,022.98)	27.68 (44.66)
Catalog only (C)	491	11.80	2.22 (.51)	1.22 (.55)	1,561.11 (5,176.72)	26.94 (61.35)
Internet only (I)	516	12.40	2.37 (1.31)	1.73 (1.19)	639.47 (622.36)	16.04 (40.74)
Brick and catalog (B + C)	252	6.05	3.98 (2.27)	3.12 (2.29)	2,204.65 (4,492.55)	50.12 (66.41)
Brick and Internet (B + I)	203	4.88	4.14 (2.25)	3.89 (2.39)	720.13 (202.61)	37.24 (82.82)
Catalog and Internet (C + I)	42	1.01	4.06 (1.58)	2.60 (1.58)	2,223.09 (6,702.44)	60.95 (68.87)
Brick, catalog, and Internet (B + C + I)	38	.91	5.71 (3.66)	3.88 (2.55)	2,379.97 (6,585.03)	72.48 (116.40)
Total	4162					
Multiple Channel User						
Yes	535	12.85	4.15 (1.77)	3.40 (2.13)	1,612.93 (6,256.16)	27.91 (41.47)
No	3627	87.15	3.05 (1.97)	3.01 (2.06)	710.79 (3,552.74)	38.39 (75.05)
Total	4162					

Summary of Planned Contrasts

	Number of Unique Purchase Occasions	Number of Categories Purchased in	Total Dollars Spent over Relationship	Number of Items Purchased
Single-channel buyer	B > C = I	B > I > C	C > B = I	B = C > I
Two-channel buyer	B + C = B + I =C + I	B + I > B + C > C + I	B + C > B + I B + C = C + I B + I = C + I	C + I > B + C > B + I
Three-channel buyer versus two-channel buyer	B + C + I > any 2	B + C + I > B + C B + C + I = B + I B + C + I > C + I	B + C + I = B + C B + C + I > B + I B + C + I = C + I	B + C + I > B + C B + C + I > B + I B + C + I = C + I

Table 2: Choice Model Variables

Variable Name	Operationalization
Price	Dollar amount paid for the product
Product category	Ten dummy variables that represent the 11 basic product categories into which all of the retailer's products are grouped
Distance	Number of miles the customer lives from the closest store
MARCOM dollars	Time-varying measure that equals the dollar amount of the direct marketing communications that were sent to each customer after his or her prior purchase and the current purchase at time t
(MARCOM dollars)^2	The square of the MARCOM dollars
Number of MARCOM	Time-varying measure that equals the number of direct communications the customer receives after his or her last purchase and before the current purchase at time t
Prior channel	Two dummy variables that indicate the prior channel from which the customer made a purchase
Purchase occasion	Time-varying measure that equals the current purchase occasion number of the customer

where

$L(H_k|i)$ = the likelihood that customer k has channel choice history H given that he or she is in segment i,

$P_j(u_{ji},\beta_{ji},X_{kt})$ = the probability that customer k chose channel j at time t given, that he or she is in segment i,

u_{ji} = the preference parameter for channel j given that the customer is in segment i,

β_{ji} = the coefficient vector for channel choice j given that a customer is in segment i, and

X_{kt} = the vector of covariates for customer k at time t.

In the case of these data, we derive the probability in Equation 1 using the parameter estimates from Step 1 of this MARCOM process. In Step 1, we also estimated a parameter that leads to the determination of the size of each latent segment. Specifically, we found that 27% of the customers likely belong to Segment 1 and that 73% likely belong to Segment 2. Using the segment size information, for each customer, we compute the probability that he or she belongs to a particular segment. Borrowing from the work of Kamakura and Russell (1989, Eq. 8), we describe this posterior probability of segment membership in Equation 2:

$$(2) \qquad P(k \in i|H_k) = \frac{L(H_k|i)f_i}{\sum_{i'} L(H_k|i')f_{i'}}$$

where

$P(k \in i|H_k)$ = the probability that customer k is in segment i given choice history H,

f_i = the probability of being in segment i, and

i' = the identifier of a latent segment

On the basis of the result of Equation 2, we assign a customer to the segment for which he or she has the highest probability of membership.

After all customers have been assigned to a segment, we profile the segments on the basis of the key variables of prior purchase behavior, demographics, and the nature of the communications between the firm and the customers. Table 4 shows the profile for these data. Note that the profile may contain variables that were not included as independent variables in the channel choice model. We use the information in Table 4 to help classify prospects in Step 5.

Step 3: Predict the Probability of Channel Choice over Time

The goal of this step is to develop a series of Markov switching matrices that reflect the probability of choosing a specific channel in the next period, given that the current channel choice is known. Using the parameter estimates from Step 1 that appear in Table 3, for each segment, we predict the probability of choice at time t at the mean value for the continuous variables (except the purchase occasion variable) and the modal value for the categorical variables (except the prior channel variable). Consistent with our definition, we set the purchase occasion variable equal to two

Tabl: 3: Logit Model Estimates

	Silgmen1 1				Segmtrtt 2			
	8rtdc V. In		OMIIIOg Venua ln. .rMte		Brkk Venua lntal'nM.•		Ca1a10Ql V.raua lnlemet•	
	Coefftclerl (Sia rd Emn)	£16aUettyaJ	eo.tftld•n (Siandlrd !Enot)	Ewtltltr	Coefflc to (Sl:andanl Enot)	e:r.llcltyc	Codlc.IMI'O (Standlrd Elror)	e.. .trc:f
lnt'ert':IGpt	–.865		.228		-6..014	''	–7.2?'1	
	(.08&)		(.083)		(.110)		(.130)	
Pwt:hM&	100	1.451''	–.!505	–.201''	UH8	.000–	2.034	.434· -
	(,008)		(.012)		(032)		(.033)	
P1iea	–.C18	-J1Z7d	–.017	–.018''	.on	000•	.052	–,189u
	(.001)		{.001)		(.003)		(.003)	
CammunicatJon dollaa	–0128	-..219''	–.0470	.081''	1.500	23..865'	-9.976	---6.885'''
	(.000)		(.000)		(.00'*)		(.007)	
(ConvnunicQtion dolal'1);1	-.0138	iH	–.0450		–.0445	ro	1n	
	(000)		(000)		(.000)		(.000)	
Nvmbof of comnaJnicationa	‹47&	2	–.048	–.039	.72-3	.000 ·	786	,343ro
	(.000)		(.032)		(.035)		(.055)	
O\$tance trom doses.\ bock	–.107	–12-iSO''	–.0001	.007	–.003	.C00-'	–.091	–2.520'''
	(.001)		(.000)		(.000)		(.016)	
PriOt channoJ Is brd	857	2.662	373e	.541''	2.275	.000' ·	-2..31	-4.706'''
	(082)		(.075)		(.059)		{.110)	
Prlof d\Mnet1:s catatog	857	2662''''	3.738	.541''	.403	-.002••	3.:243	2.839'''
	(083)		(.075)		{;089)		(.095)	

Table 3: Continued

	Segment 1				Segment 2			
	Brick Versus Internet[a]		Catalog Versus Internet[a]		Brick Versus Internet[a]		Catalog Versus Internet[a]	
	Coefficient[b] (Standard Error)	Elasticity[c]	Coefficient[b] (Standard Error)	Elasticity[c]	Coefficient[b] (Standard Error)	Elasticity[c]	Coefficient[b] (Standard Error)	Elasticity[c]
Product Categories								
Category 2	−.875 (.066)	.585**	−1.709 (.060)	−.249**	1.660 (.057)	.000**	.804 (.099)	−.856**
Category 3	−1.404 (.059)	−.382**	−1.195 (.056)	−.173**	.599 (.036)	.000**	−.408 (.091)	−1.007**
Category 4	−1.049 (.069)	−.637**	−.482 (.064)	−.070**	2.295 (.079)	.000**	3.062 (.116)	.767**
Category 5	−1.288 (.098)	.584**	−2.191 (.102)	−.319**	28.304 (74.200)		26.204 (67.420)	
Category 6	−.935 (.051)	.000**	−1.093 (.048)	−.159**	2.303 (.183)	.000**	7.742 .204	5.439**
Category 7	−.935 (.051)	.000**	−1.093 (.048)	−.159**	20.099 (125.249)		16.447 (25.250)	
Category 8	−.934 (.051)	−.001**	−1.092 (.048)	−.159**	1.064 (.069)	.000**	−1.049 (.141)	−2.113**
Category 9	−.933 (.052)	.000**	−1.091 (.048)	−.159**	1.162 (.065)	.000**	.708 (.187)	−.454*
Category 10	−.935 (.053)	.000**	−1.093 (.049)	−.159**	11.271 (10.225)		−81.580 (81.655)	
Category 11	−.935 (.051)	−.001**	−1.093 (.048)	−.159**	.172 (.059)	.000*	−.912 (.187)	−1.084**
Segment Size	.27				.73			

*$p < .05$.
**$p < .0001$.
[a]To determine the coefficient for choosing the brick store versus the catalog, we subtracted the coefficient for choosing the catalog versus the Internet from the coefficient for brick versus the Internet.
[b]A positive coefficient means that a customer is more likely to choose channel j than the base channel. The base channel is the Internet in this case.
[c]We computed elasticities at the mean value of the continuous variables and the modal value of the categorical variables.

for the first repeat purchase occasion. For subsequent repeat purchase occasions, we increase the variable by one accordingly. The parameter that we use in the prediction for the prior channel choice depends on which row of the matrix we are predicting. For these data, the channel choices for four periods into the future appear in Table 5. We observe that there are two distinct segments in terms of channel use: Segment 1 frequents the catalog and/or Internet, and Segment 2 is channel loyal and frequents the bricks-and-mortar store. Given the degree of loyalty, it is not surprising that the majority of Segment 2 customers made their first purchase from this same channel. This pattern of channel loyalty is consistent with Dholakia, Zhao, and Dholakia's (2005) findings.

Step 4: Develop a Segment-Specific Communications Strategy

Developing a segmented MARCOM strategy begins with some basic questions. Which customer types generate the most value to the firm? From Table 1, we confirm that multichannel buyers generate more revenue for the firm, purchase more items, purchase in more categories, and purchase more frequently than do single-channel buyers. More specifically, we learn that multichannel customers who use catalogs tend to generate more revenue than multichannel customers who do not use catalogs. We even find that a dual-channel customer is equally as valuable as a three-channel customer in terms of revenue, as long as the dual-channel customer buys from the catalog. Although this information does not imply a causal relationship between channel choice and purchase behavior, it provides a guide that we can use to develop a communication strategy for influencing channel choice.

Table 4: Segment Profiles

	Segment 1		Segment 2	
	Mean	Standard Deviation	Mean	Standard Deviation
Relationship duration (days)	138.39	166.48	137.70	152.48
Number of unique purchase occasions	2.73	1.22	3.73	2.40
Total revenues generated throughout entire relationship	850.29	434.80	727.05	3825.85
Total revenues from first purchase	106.06	236.72	90.36	161.97
Total number of items purchased throughout entire relationship	106.40	116.96	118.15	148.87
Distance from closest bricks-and-mortar store (miles)	121.62	158.78	28.66	21.21
Number of channel switches	.30	.68	.19	.68
Days elapsed between channel switches	118.84	144.49	103.51	115.62
Total amount of direct mail communication dollars	2.60	1.53	2.64	1.45
Total number of direct mail communications	5.29	5.22	5.43	5.04
Percentage of Customers with Channel of First Purchase				
Bricks-and-mortar	34.53		75.05	
Catalog	46.64		7.94	
Internet	18.83		17.01	
Percentage of Customer Purchases by Category				
Category 1	47.91		49.76	
Category 2	8.84		9.45	
Category 3	12.49		12.42	
Category 4	8.65		6.81	
Category 5	2.10		2.55	
Category 6	3.83		1.10	
Category 7	2.26		2.42	
Category 8	4.29		4.46	
Category 9	2.01		2.93	
Category 10	4.61		5.85	
Category 11	3.00		2.26	
Percentage of Customers Making up to N Repeat Purchases				
1	58.12		21.33	
2	26.22		57.44	
3	7.79		13.21	
4	4.44		2.90	
5	1.89		2.77	
6	.51		.85	
7	.36		1.01	
8	.29		.35	
9	.36		.13	
10+	.00		.00	
Percentage of Customers in Segment	27		73	

What are the customer's intrinsic channel preferences and tendencies? The answer to this question comes primarily from the Markov matrices we developed in Step 3 (see Table 5). These matrices reveal that Segment 1 is primarily a catalog segment, as long as the prior purchase was not from the Internet. If the customer's prior channel was the bricks-and-mortar store, the probability is high that his or her first and second repeat purchases will be from the catalog. If the customer's prior purchase was over the Internet, there is some migration toward the catalog in the early stages of the life cycle, but this diminishes as the repeat purchase occasions increase. In general, Segment 1 customers could be labeled as those who migrate toward remote channel.

In contrast, Segment 2 customers will most likely stay in the bricks-and-mortar channel or switch to it. Over time, the data predict that Segment 2 customers will not choose to buy from the Internet. Given the infrequent switching that occurs in this segment (see the profile in Table 4), it could also be concluded from the trajectory of the matrices that a small group of customers in Segment 2 who repeat buy from the catalog will likely be those whose first purchase was from this same channel.

Thus, consistent with the segment profiles, these forecasts indicate that there is a significant amount of channel stickiness for buying in both segments. This is an important finding for several reasons. First, it suggests that for these data, the channel of first purchase has a high probability of being the channel choice for subsequent purchases. Second, this finding leads to the next set of questions.

To what extent do customers respond to MARCOM, and what is the nature of their response? On the basis of the elasticities in Table 3, we conclude that the number of communications is a key factor that is associated with the choice of the bricks-and-mortar store over the Internet but not for the choice between the catalog and the Internet in Segment 1 (see Table 3). The prior channel also seems to play a significant role in subsequent channel choices for Segment 1. Specifically, the lack of prior experience on the Internet is related to the decision not to choose the Internet for subsequent purchases. For Segment 2, the elasticities (see Table 3) indicate that the MARCOM dollar expenditures appear to drive channel choice the most. Increasing the number of communications in Segment 2 is associated with the choice of the Internet over the store or the catalog.

Table 5: Forecasted Channel Switching Probabilities

	Transition Matrices					
	Segment 1: Next Period Channel			Segment 2: Next Period Channel		
	Bricks-and-Mortar	Catalog	Internet	Bricks-and-Mortar	Catalog	Internet
First Repeat: Current Channel						
Bricks-and-mortar	.000	.886	.114	.992	.001	.007
Catalog	.000	.886	.114	.737	.231	.032
Internet	.000	.198	.802	.888	.049	.063
Second Repeat: Current Channel						
Bricks-and-mortar	.000	.828	.171	1.000	.000	.000
Catalog	.000	.828	.171	.741	.259	.000
Internet	.000	.132	.868	.941	.059	.010
Third Repeat: Current Channel						
Bricks-and-mortar	.001	.751	.248	.999	.001	.000
Catalog	.001	.751	.248	.721	.279	.001
Internet	.000	.085	.915	.942	.056	.001
Fourth Repeat: Current Channel						
Bricks-and-mortar	.001	.655	.344	.999	.001	.000
Catalog	.001	.655	.344	.700	.300	.000
Internet	.000	.054	.946	.940	.060	.000

Given the significance of both the linear and the quadratic terms in both segments, we show the effect of MARCOM dollar expenditures graphically in Figure 1. This figure shows the change in the log odds of channel choice with respect to MARCOM dollar expenditures. To compute the log-odds ratio, we assume that the customer is making his or her first repeat purchase and that all other continuous variables are fixed at their means and the categorical variables are fixed at their modes.[4] For Segment 1, we find that the log-odds ratio is negative and decreasing for the bricks- and-mortar store and positive and decreasing for the catalog channel. This means that an increase in MARCOM spending decreases the chance of the customer's choosing the bricks-and-mortar channel over the Internet. In addition, overall, MARCOM expenditures enhance the chances of the customer's choosing the catalog over the Internet, but this probability decreases as MARCOM expenditures increase. Thus, within the range of these data, we conclude that, in general, increasing MARCOM expenditures directed at the average Segment 1 customer motivates him or her to make the first repeat purchase from the catalog.

For Segment 2, we find that the log-odds ratio is positive and increasing for the bricks-and-mortar store and negative and decreasing for the catalog channel.[5] This means that an increase in MARCOM spending increases the chance of choosing the bricks-and-mortar channel over the Internet and decreases the chance of choosing the catalog over the Internet. On the basis of the magnitude and sign of the log-odds ratios, we conclude that higher levels of communication spending directed toward Segment 2 primarily drives the average Segment 2 customer to make his or her first repeat purchase in the bricks-and-mortar store.

Synthesizing all of this information leads to an initial strategy for communicating with customers. The objective of this strategy is to encourage channel choice behavior that is consistent with the firm's highest-value multichannel customers. Because we are interested in driving net marketing contribution, we ignore the fixed costs that are associated with operating the channels when developing this strategy. Given this objective and the responses to the critical questions, we conclude that though Segment 1 customers are more likely to switch channels than Segment 2 customers, MARCOM plays a more limited role in this decision. The prior purchase channel is a vital if not better predictor of subsequent channel choice.

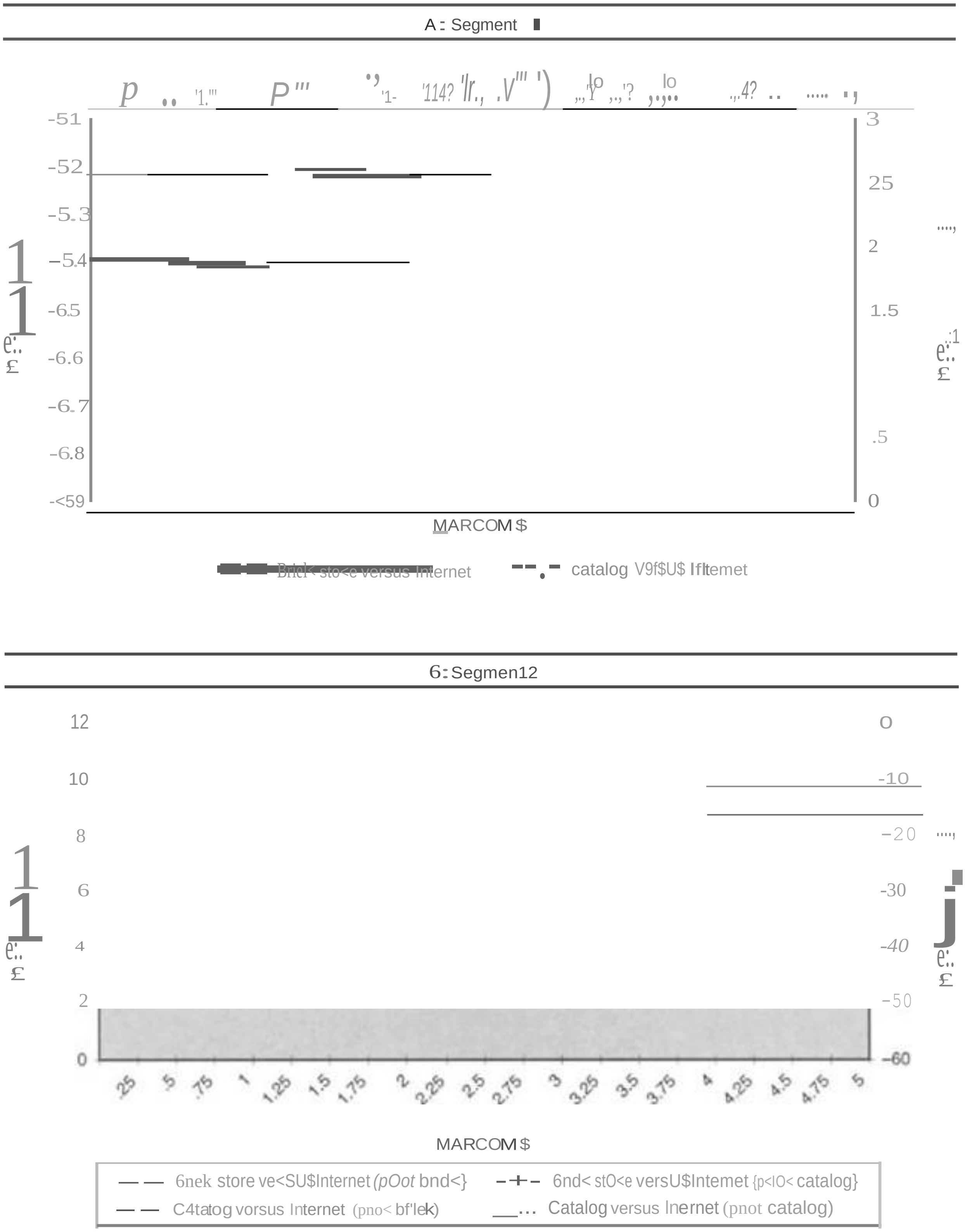
Many Segment 1 customers (46%) made their first purchase through the catalog. For customers in this portion of the segment, the retailer should attempt to maintain its strong affiliation with the catalog channel and encourage customers to have a multichannel relationship that likely would include the Internet. The rationale for this is that historical data indicate that this type of multichannel customer is desirable because of his or her tendency to be one of the highest dual-channel revenue generators, have a high purchase volume, and purchase at least as often as any other dual-channel buyer (see Table 1). Given that the number of communications does not have a significant effect on the choice between the catalog and the Internet channels, the firm should increase the MARCOM expenditure but direct the increase toward enhancing the quality of the communications, not the number of communications. An opportunity for improving the quality in a way that motivates the use of the Internet would be to feature products that are not in the base product category (i.e., Category 1; see Table 3).

For the 54% of Segment 1 who began a relationship through a channel other than the catalog, the goal is to drive these customers toward the catalog. The odds are against the choice of the bricks-and-mortar store over the catalog, but they improve slightly as the MARCOM dollars increase from $.25 to $5.00. Similarly, the odds are against the choice of the Internet over the catalog, but they improve as MARCOM expenditures increase from $.25 to $5.00. In addition, we know that increasing the number of

4 Given that in the data the average number of unique purchases is between two and three, we determined that it was best to assume that the customer was making his or her first repeat purchase.

5 Although we do not show this in Figure 1, we also used the parameter estimates to assess the choice between the physical bricks-and-mortar store and the catalog.

Figure 1: Effect of MARCOM Expenditure on hannel Choice



Note: The [illegible] of the vertical y-axes should be conceptualized as the odd of choosing the specific channel versus the Internet.

communications drives this group of customers away from the catalog (see Table 3). On the basis of all this information, we conclude that the best way to drive these customers toward the catalog is to limit the amount of MARCOM spending and the number of contacts.

Although both segments exhibit channel loyalty, Segment 2 customers exhibit even higher levels of channel loyalty. This makes the channel of first purchase an even more important predictor of future choices. This also suggests that encouraging a person in Segment 2 to use all three channels is extremely difficult. Given that Segment 2 is channel loyal and that catalog-only customers generate the most revenue of all the single-channel customers (see Table 1), with limited MARCOM resources, it is better to focus a channel migration initiative on customers in Segment 2 whose first purchase was not through a catalog. Using the data in Table 1, we determine that the most desirable migration profiles for this subgroup are for first-time bricks-and-mortar store customers to become store and catalog customers and for first-time Internet customers to migrate to either of the other two channels. In addition, the estimates and elasticities from Table 3 suggest that for this subgroup, both the amount and the number of MARCOM expenditures are instrumental in achieving these migration objectives.

In Table 6, we summarize the specific MARCOM goals, a tactical plan, and the expected outcome of that plan for both segments. We derived the financial outcomes that appear in Table 6 using a segment-specific comparison of the revenue generation from a single-channel customer becoming a dual-channel customer (e.g., a catalog-only customer becoming a catalog and Internet customer in Segment 1). On the basis of these expected financial outcomes, we conclude that the potential payout may be greater if the firm focuses its incremental MARCOM resources on customers whose first channel was not the catalog.

Although the financial projections of the initial MARCOM strategy appear promising, it is critical that the firm asks one final question before moving forward: What are the costs associated with the different MARCOM activities? This information was not available for these data. However, before adopting any MARCOM plan, the expected payout must be compared with the expected costs that are used to generate that payout.

Table 6:
Initial MARCOM Strategy

	Segment 1	Segment 2
First Purchase Channel Is Catalog	•Migration objective: Maintain catalog affiliation, and migrate to the Internet. •Tactical plan: Focus on increasing the MARCOM dollar expenditures, and leverage this increase to enhance the quality of MARCOM. •Expected plan outcome: Customer profile: C → C + I Change in revenue: 118%	•Migration objective: Given limited resources, accept these customers as single-channel customers. •Tactical plan: Maintain status quo MARCOM. •Expected plan outcome: Customer profiles: C → C Change in revenue: 0%
First Purchase Channel Is Not Catalog	•Migration objective: Encourage migration to catalog channel. •Tactical plan: Limit both the amount and the number of expenditures, and leverage existing expenditures to enhance the quality of existing MARCOM. •Expected plan outcome: Customer profiles: B → B + C, or I → I + C Change in revenue: 138%–199%	•Migration objective: Drive customers to become dual-channel customers, specifically bricks-and-mortar store and catalog or Internet plus any other channel. •Tactical plan: Focus on increasing the MARCOM dollar expenditures, and leverage this increase to enhance both the quality of communications and the number of MARCOM activities. •Expected plan outcome: Customer profiles: B → B + C, I → I + B, or I → I + C Change in revenue: 14%–250%

Step 5: Classify Prospects and First-Time Customers into Existing Segments

A significant benefit of this MARCOM process is using it to classify prospects and first-time customers. The earlier a firm can accurately classify a customer, the more efficiently and effectively it can leverage its MARCOM plan. In this step, we classify customers before their first purchase and update this classification after the first and second purchases. We iterate the classification process three times to demonstrate the accuracy and degree of improvement in classification that comes with added customer interaction data.

In this step, we used a tree approach for classifying customers (Breiman et al. 1984). Specifically, we used both CHAID and CART and compared the results in terms of predictive accuracy. The segment profile in Table 4 provides the best guide for variables that can be used to classify prospects and customers. The usefulness of the profile depends on the amount of information the firm has about customer characteristics, early purchase behavior, and the distinctiveness of that information across segments. In these data, the most distinguishing pieces of information are the distance the customer lives from the bricks-and-mortar store and the channel of first purchase. The channel choice estimates from Step 1 of this process confirm that distance is a significant driver of channel choice. Given the channel migration patterns that we uncovered in Step 3, we can conclude that the sequence of channel choices from the first to the second purchase might also aid in classification.

Thus, using only the distance from the closest store, we estimated a classification tree using a training sample of 2081 customers, and we tested it on a different set of 2081 customers. We found that, in general, the CART and CHAID results were similar, but the CART procedure was slightly more accurate with respect to prediction in Segment 1, the segment that generates the most revenues. From the CART results, we found that the overall risk of misclassification using only the distance measure was 21.39%, which we computed by dividing the total number of incorrectly classified customers in the test sample by the total number of customers in the test sample. We accurately classified 52.15% of the Segment 1 customers and 87.69% of the Segment 2 customers; we computed this by dividing the number of customers that CART predicted to be in a segment by the number that are truly in that segment. Using this classification, an initial contact strategy consistent with segment behavior and firm goals would be to select Segment 1 customers and send them a more extensive catalog that emphasizes the breadth of the product line and informs them of the Internet channel. In contrast, the initial contact for potential Segment 2 customers would be a communication piece (e.g., a postcard) that encourages them to make a purchase in a bricks-and-mortar store.

Although the channel of first purchase is not a perfect discriminator of segment membership, it can help classify a fraction of the first-time buyers. This is because a comparison of the segment profiles in Table 4 suggests that a customer whose first purchase is from the catalog is most likely a Segment 1 customer. However, given that a sizable number of customers in both segments make their first purchase in a store or on the Internet (see Table 4), it is difficult to assign these customers to segments. Using both the distance and the channel of first purchase, the CART procedure accurately classified more Segment 1 customers (74.55%) and fewer Segment 2 customers (78.46%). Because Segment 2 is significantly larger than Segment 1, the overall risk of misclassification increases to 23.49%.

On the third iteration of the CART procedure, we used distance, channel of first purchase, and channel of second purchase to classify customers. The motivation for adding this third variable was based on the degree of channel stickiness in Segment 2 and the migration toward the catalog in Segment 1 (see Table 5). The results from this iteration showed that there was only a 12.05% risk of customer misclassification. We accurately classified 74.01 % of the Segment 1 customers and 92.74% of the Segment 2 customers. Thus, as the firm gains information, the classification improves. However, using only information that is known before the first purchase (i.e., distance from the store), for these data, the CART procedure predicted segment membership with approximately 80% accuracy.

Step 6: Update Segment Affiliation

This step reiterates Steps 1-4 on existing customers. The timing of when to implement Step 6 depends on the dynamics in a firm's market and the degree to which customer behavior changes over time.

SUMMARY, LIMITATIONS AND FURTHER RESEARCH

Although the insights derived from this analysis are specific to these data, the process that we describe is generalizable. Our application demonstrates that analyses can be conducted to enhance the targeting and management of customers in a multichannel context, even if a firm is just beginning to develop an enterprise database. A key benefit of this process is that it leads to insights that can be used to classify prospects into segments quickly. As MARCOM tactics become more sophisticated and costly, early and accurate customer identification is even more critical, and this process becomes more valuable.

A limitation of the inferences that we derive from using the process is that our data do not account for the nature or content of the MARCOM. For example, the device might be a specialty coupon, such as a buy-one-get-one-free pro- motion, or it may be a postcard that focuses on a particular category. Given our

data, we simply examine the dollar expenditures and the number of MARCOM activities. An avenue for further research is to apply this process to more specific data on the communication devices themselves.

Another area for further research is to gather and incorporate shopping-level data with the purchase data to track customers through all interactions with the retailer (Nunes and Cespedes 2003). The data can be included in the choice model specification as we describe in Step 1. In addition, there is a need for theoretical models of customer buying behavior across multiple channels. Schoenbachler and Gordon (2002) introduce one such model, and researchers such as Kumar and Venkatesan (2005) have tested some of its components and found promising preliminary results for multichannel shopping behavior. The current research is a first step toward helping retailers understand the value and use of enterprise data. Although we demonstrate how MARCOM can benefit from the investments made in developing enterprise data, research in other areas, such as inventory management (e.g., Bendoly et al. 2005), could also benefit from the data that comes from channel integration

REFERENCES

Aberdeen Group (2004), "The Integrated Multichannel Benchmark Study: Gaining Competitive Advantage by Fulfilling Multichannel Demand," (June), (accessed October 5, 2004), [available at http://www.aberdeen.com/summary/report/benchmark/multichannel_selling.asp

Anderson, T.W. and Leo A. Goodman (1957), "Statistical Inference About Markov Chains," The Annals of Mathematical Statistics, 28 (1), 89-110.

Ansari, Asim, Carl Mela, and Scott A. Neslin (2005), "Customer Channel Migration," working paper, Tuck School of Business, Dartmouth College

Balasubramanian, Sridhar, Rajagopal Raghunathan, and Vijay Mahajan (2005), "Consumers in a Multichannel Environment: Product Utility, Process Utility, and Channel Choice," Journal of Interactive Marketing, 19 (2), 12-30.

Bendoly, Elliot, James Blocher, Kurt M. Bretthauer, Shanker Krishnan, and M.A. Venkataramanan (2005), "Online/In-Store Integration and Customer Retention," Journal of Service Research, 7 (4), 313-28.

Berger, Paul D., Ruth Bolton, Douglas Bowman, Elten Briggs, V. Kumar, A. Parasuraman, and Creed Terry (2002), "Marketing Actions and the Value of Customer Assets: A Framework for Customer Asset Management," Journal of Service Research, 5 (1), 39-54.

Boulding, William, Ajay Kalra, and Richard Staelin (1999), "The Quality Double Whammy," Marketing Science, 18 (4), 463-84.

Breiman, L., J.H. Friedman, R.A. Olshen, and C.J. Stone (1984), Classification and Regression Trees. Belmont, CA: Wadsworth International Group.

Brynjolfsson, Erik and Michael D. Smith (2000), "Frictionless Commerce? A Comparison of Internet and Conventional Retailers," Management Science, 46 (4), 563-85.

Cleary, M. (2000), "The Promise of Multichannel Retailing," Inter@ctive Week, (October 9), 50.

Dholakia, Ruby Roy, Miao Zhao, and Nikhilesh Dholakia (2005), "Multichannel Retailing: A Case Study of Early Experiences," Journal of Interactive Marketing, 19 (2), 63-74.

Eagly, Alice H. and Shelly Chaiken (1993), The Psychology of Attitudes. Fort Worth, TX: Harcourt Brace Jovanovich.

Forster, Stacy (2004), "When One Hand Doesn't ...," The Wall Street Journal, (March 22), R3.

Fox, Edward J., Alan L. Montgomery, and Leonard M. Lodish (2004), "Consumer Shopping and Spending Across Retail Formats," The Journal of Business, 77 (2), 525-560.

Kamakura, Wagner and Gary J. Russell (1989), "A Probabilistic Choice Model for Market Segmentation and Elasticity Structure," Journal of Marketing Research, 26 (November), 379-90.

Kumar, V. and Rajkumar Venkatesan (2005), "Who Are the Multi-channel Shoppers and How Do They Perform? Correlates of Multichannel Shopping Behavior," Journal of Interactive Marketing, 19 (2), 63-74.

Montgomery, David B. (1969), "A Stochastic Response Model with Applications to Brand Choice," Management Science, 15 (7), 323-37.

Nunes, Paul F. and Frank V. Cespedes (2003), "The Customer Has Escaped," Harvard Business Review, 81 (November), 96-105.

Reinartz, Werner, Jacquelyn S. Thomas, and V. Kumar (2005), "Balancing Acquisition and Retention Resources to Maximize Customer Profitability," Journal of Marketing, 69 (January), 63-79.

Schoenbachler, D.D. and G.L. Gordon (2002), "Multi-Channel Shopping: Understanding What Drives Channel Choice," Journal of Consumer Marketing, 19 (1), 42-53.

Shankar, Venkatesh, Arvind Rangaswamy, and Michael Pusateri (2001), "The Online Medium and Price Sensitivity," working paper, Mays Business School, Texas A&M University.

Styan, George P.H. and Harry Smith Jr. (1964), "Markov Chains Applied to Marketing," Journal of Marketing Research, 1 (February), 50-55.

Tellis, Gerard J. (2003), "Econometric Estimation of Media Effectiveness and the Optimal Allocation of Media Expenditures," in Measuring and Allocating Marcom Budgets: Seven Expert Points of View. Cambridge, MA: Marketing Science Institute, 45-52.

Thomas, Jacquelyn S., Robert C. Blattberg, and Edward Fox (2004), "Recapturing Lost Customers," Journal of Marketing Research, 41 (February), 31-45.

Young, Sam J. (2001), "Make the Most of Every Channel," Catalog Age, 18 (March), 61-62.

Acknowledgement: Reprinted with permission from Journal of Marketing, published by the American Marketing Association, Thomas, J.S. & Sullivan, U.Y., 2005, Volume 69(4): 239 – 251.