Marketing Science Institute Special Report 07-207

Optimal Allocation of Marketing Efforts by Customer-Channel Segment

Tarun Kushwaha and Venkatesh Shankar

Optimal Allocation of Marketing Efforts by Customer-Channel Segment

Tarun Kushwaha
Assistant Professor of Marketing
Kenan-Flagler Business School
University of North Carolina
CB# 3490 Chapel Hill, NC 27599-3490
Ph: (919) 962-8746, Fax: (919) 962-7186
E-mail: tarun_kushwaha@unc.edu

Venkatesh Shankar
Professor of Marketing and Coleman Chair in Marketing
Department of Marketing
Mays Business School
Texas A&M University
4112 TAMU
College Station, TX 77843-4112
Phone: (979) 845-3246, Fax: (979) 862-2811
E-mail: VShankar@mays.tamu.edu

November 2007

*We thank a large manufacturer and marketer of shoes and accessories for providing the data, Len Berry for his help in acquiring the data, the Levy and Weitz Retailing SIG dissertation award for doctoral support, and participants at the Marketing Science Conference 2006, at research seminars at the University of North Carolina, Syracuse University, University of Illinois at Urbana-Champaign, Georgia Institute of Technology, University of South Carolina, and McGill University and the MSI review team for valuable comments. This work is part of the first author's dissertation.

Optimal Allocation of Marketing Efforts by Customer-Channel Segment

ABSTRACT

As firms increasingly offer their products through multiple channels such as store, the Web and catalog or direct mail and as more consumers buy them through different channels, the allocation of marketing efforts targeted at customers across channels is becoming a critical issue for many marketers. We propose an approach and model for *optimal allocation* of marketing efforts to each customer-channel segment. We first develop marketing response models for each component of firm profit, *purchase frequency*, *purchase quantity*, *product return propensity*, and *contribution margin*. We model purchase frequency using the extended Beta Geometric/Negative Binomial Distribution model, purchase quantity and product return propensity using the Conditional Negative Binomial Distribution model, and contribution margin using the Gamma-Gamma model. The optimal marketing effort allocation to each customer-channel segment is a function of the model parameters for that segment. We estimate the models using customer level purchase, cost, and promotional data from a large marketer of shoes and apparel accessories across multiple channels, namely, the catalog, the store, and the Web. We solve the optimization model using simulation. The optimization model can be implemented in Excel. The results show that consumer response to firm marketing efforts varies significantly across the customer-channel segments for the different profit components, purchase frequency, purchase quantity and contribution margin. Using a holdout sample analysis, we show that firm profits can be substantially improved by optimally reallocating marketing efforts across the different customer-channel segments. In the revised allocation, the multichannel segment exhibits the highest percentage growth in budget and profit, highlighting the high profit potential of the multichannel segment.

Keywords: Multichannel management, Resource Allocation, Statistical Models, Optimization, Segmentation.

Introduction

The allocation of marketing resources across various marketing instruments and customer[1] segments is a topic of immense interest to marketing academics and practitioners alike. Studies in marketing have examined allocation of marketing resources across sales force (Rangaswamy, Sinha, and Zoltners 1990), markets (Mantrala, Sinha, and Zoltners 1992), media (Naik, Mantrala, and Sawyer 1998), mailing campaigns (Elsner, Krafft, and Huchzermeier 2004), acquisition and retention efforts (Reinartz, Thomas, and Kumar 2005), and customers (Venkatesan and Kumar 2004).

Organizations deploy marketing resources through multiple channels such as physical store, the Web, and catalog. Customers may choose one or more of these channels for their purchases and can be segmented based on their channel choice, leading to the identification of customer-channel segments (Kumar and Venkatesan 2005; Thomas and Sullivan 2005). In particular, the multichannel segment constitutes a valuable customer segment for marketers (e.g., Doubleclick 2004; Kumar and Venkatesan 2005). By knowing how different customer-channel segments respond to marketing efforts, managers can better allocate their marketing resources across these segments (Neslin and Shankar 2007).

The need to address resource allocation decisions at customer-channel segment level has been raised by practitioners and academia alike. Neslin et al. (2006) emphasize the need to develop models for the allocation of marketing resources across channels.[2] Libai, Narayandas, and Humby (2002) argue that channel-segment based approach to resource allocation can bring significant improvement in profitability. Similarly, studies by IBM and McKinsey & Company call for developing resource allocation metrics across channels (Achabal et al. 2005; Myers, Pickergill, and van Metre 2004). The primary advantage of allocating marketing resources at the customer-channel segment level is that it provides firms with the ability to leverage channel usage as a segmentation tool and utilize marketing instruments in different channels with differing intensities.

[1] For expositional use, we use the terms, consumer and customer, interchangeably throughout the paper.
[2] We do not address allocation across channels, which may involve the fixed costs of set up and maintenance.

In this paper, we address three managerially important research questions relating to the allocation of marketing efforts across customer-channel segments. First, how much marketing efforts should a firm expend for each customer-channel segment? Second, can a firm improve its profitability by incorporating multichannel shopping behavior in its resource allocation decisions? Third, can we decompose the responsiveness of profits to marketing efforts in each customer-channel segment into the responsiveness of purchase frequency, purchase quantity and contribution margin to marketing efforts?

We develop models to estimate the responsiveness of different customer-channel segments to marketing efforts. Based on these response parameters, we develop a resource allocation model to optimize the allocation of marketing resources across customer-channel segments. We decompose customer profits into profits due to purchase frequency (number of orders), purchase quantity per order, product return propensity, and contribution margin per item by developing models for each of these components. For purchase frequency, we use an extended Beta Geometric/Negative Binomial Distribution model that includes the effect of marketing covariates on purchase frequency. We model purchase quantity and product return propensity, using Conditional Negative Binomial Distribution models. We model contribution margin per item using a Gamma-Gamma model. We estimate these models using data using customer level purchase, cost, and promotional data from a large marketer of shoes and apparel accessories across multiple channels. We solve the optimization model using simulations. Although the percentage of multichannel shoppers in our dataset is somewhat small, our optimization model is general and can be implemented by any firm regardless of the relative composition of channel segments, using Excel software.

Our modeling approach extends related existing research in many ways. First, to our knowledge, ours is the first to offer a rigorous, yet practical approach to the allocation of marketing efforts at the customer-channel segment level. Prior research has either developed descriptive models of response to marketing efforts in a multichannel context (e.g., Venkatesan, Kumar, and Ravishanker 2007) or offered contact optimization at the customer level (e.g., Venkatesan and Kumar 2004). Because allocation of marketing efforts at a customer-channel segment level is becoming more practical and cost-effective for firms (e.g., *J.C. Penney 10K Statement* 2006), our optimization approach at the customer-channel segment level offers an important and useful tool to marketers. Second, existing customer-level approaches (e.g., Fader, Hardie, and Lee 2005a, b; Kumar and Venkatesan 2005) decompose model contribution margin at an *order* level by splitting it into two elements: purchase frequency and contribution margin per order. Our approach extends these approaches by modeling contribution margin at an *item* level by decomposing it into four components: purchase frequency, purchase quantity per order, product return per item, and contribution margin per item. Such an approach enables us to differentiate between order size effect and up-selling effect. These effects may be different and need to be captured separately to derive an optimal customer-channel marketing effort allocation model. These effects, however, have not been accounted for by prior studies. Third, prior stochastic models of customer purchase behavior in an interactive marketing context (e.g., Fader, Hardie, and Lee 2005a, b) do not include marketing covariates and hence do not offer optimal marketing allocation results. Our models include the effects of marketing efforts and offer an approach to optimization of marketing resources. Fourth, previous resource allocation models tend to ignore customers' product return propensities, leading to potentially inflated firm profit. Anecdotal evidence suggests that product returns can range from 10% to 25% of the orders, depending on the product category (Fenvessy 1992; Hess and Mayhew 1997). By incorporating product return propensity into our model, we use a more appropriate metric for computing firm profit.

Our results show that the responsiveness of different customer-channel segments to marketing efforts varies substantially across the different components of profit and across different customer-channel segments. Using a holdout sample analysis, we also show that a firm can improve its total profits by optimally allocating marketing resources to these customer-channel segments based on the heterogeneous response behavior of these segments to the firm's marketing efforts.

The rest of this paper is organized as follows. In the next section, we develop the conceptual model. In the third section, we develop a disaggregated model for firm profit, which we decompose into elements each with a model. In the fourth section, we discuss the customer level transaction data we use for estimating the models. In the fifth section, we discuss model estimation and the results. We also evaluate the predictive validity of each model for each customer-channel segment. We conclude by discussing the managerial implications and summarizing the work.

Conceptual Development

We address the resource allocation problem across customer segments derived from channel choice behavior. We base the resource allocation decisions on the predicted future profits of the firm. The primary advantage of developing such a forward looking resource allocation metric at the customer-channel segment level is the ability to use channel as a segmentation tool. We present the conceptual framework in Figure 1. In this framework, a firm's marketing instruments (marketing mailer, promotional discount, and price) influence purchase behavior (purchase frequency, purchase quantity per order, and contribution margin per item). We anticipate that the effect of marketing instruments on customer's responsiveness will be different across the different components of profit and across the customer-channel segments. The primary thesis of this work is that (1) segmenting by channel choice is an efficient way of segmenting customers; (2) different customer-channel segments respond differently to a firm's marketing efforts; and (3) firms can improve their profits by optimally allocating marketing resources to these customer-channel segments based on the responsiveness of these segments to the firm's marketing efforts.

< Insert Figure 1 here >

There are important theoretical reasons to expect differential responsiveness of different customer-channel segments to marketing efforts. First, different channels offer different opportunities for customers to interact with the firm and these interactions could lead to different responses to marketing efforts (Berger et al. 2002; Neslin and Shankar 2007; Villaneueva, Yoo, and Hanssens 2008). For example, a customer who shops only at a store may respond to a marketing mailer by buying a wide assortment and large quantities of the items of selected products to amortize the cost of a trip to the store (Bhatnagar and Ratchford 2004; Messinger and Narasimhan 1997). In contrast, a customer who buys only on the Web, may not buy a large quantity in the first order before verifying the correctness of choice of the item after delivery (Bart et al. 2005). Second, different channels provide different levels of involvement for customers, leading to different allocation of attention resources for information processing and thereby differential levels of responsiveness to marketing activities (Assael 1998). Third, each channel calls for a unique level of cognitive effort on the part of customer to be able to react to marketing messages (Balasubramanian, Raghunathan, and Mahajan 2005). In particular, a customer's response to marketing communication on the Web may be quite different from those in other channels (Shankar and Hollinger 2007). Fourth, the relationships among communication intensity, commitment and trust are different for different channels (Morgan and Hunt 1994).

Model Development

Resource Allocation Model

The firm's objective function, total profits, is the sum of profits generated by each customer-channel segment of the firm. Let Π be the total profits of a firm over a given time horizon and Π_k be the profit of the k[th] customer-channel segment over the same horizon. Because the time horizon for each segment is the same, for ease of exposition, we drop the time subscript from the equations. The total profit is given by:

$$\Pi = \sum_{k=1}^{K+1} \Pi_k (m_{ik}) \tag{1}$$

where m_{ik} is the number of marketing mailers sent to customer i from customer-channel segment k and is the resource allocation variable. A firm that markets through K channels has K single-channel segments and one customer-channel segment of 'multichannel' users.[3] We want to determine the optimal levels of marketing efforts that would maximize Π_k for each customer-channel segment and Π for the firm across the customer-channel segments. The optimization equation is given by:

[3] For parsimony, we consider only one broad multichannel segment although the analysis could be extended to multichannel sub-segments such as store and catalog only sub-segment and catalog and Web only sub-segment.

$$\max_{m_k} \sum_{k=1}^{K+1}\sum_{i=1}^{n_k}[(IPO_{ik}(m_{ik})-IRPO_{ik})\times NO_{ik}(m_{ik})\times \overline{CM}_{ik}(m_{ik})-c_m m_{ik}] \tag{2}$$

$$s.t. \quad m_{ik} \geq 0$$

where, purchase frequency (NO_{ik}), purchase quantity per order (IPO_{ik}), and gross contribution margin ($\overline{CM}_{ik}$) of customer i from customer-channel segment k are endogenous variables that are functions of marketing efforts (m_{ik}) expended by the firm toward that customer. $IRPO_{ik}$ is product returns per order for customer i from customer-channel segment k and c_m and n_k are the unit cost of marketing and size of customer-channel segment k, respectively. We develop separate models for these endogenous variables. The detailed derivation of the profit equation appears in Appendix A.

We account for the fact that customers' channel preferences evolve over a period of time by subsequently estimating customer transition probabilities across channel segments over time and by incorporating this transition matrix in our resource allocation model. We elaborate on this issue in the results and robustness checks (Appendix C) sections of the paper.

Consistent with prior studies (e.g., Fader, Hardie, and Lee 2005b), we assume that the purchase frequency of a customer is independent of her contribution margin per item.[4] However, we do not assume that the customer's purchase quantity per order and returns per order are independent of her purchase frequency. It is reasonable to assume that customers with a greater purchase quantity per order may buy less frequently than those with a smaller purchase quantity per order. Similarly, the probability of returning an item is high when purchase frequency is high. Thus, we condition the predicted purchase quantity per order and predicted returns per order model on the predicted purchase frequency.

Marketing cost is a linear function of the amount of marketing efforts undertaken by a firm, consistent with Mantrala, Sinha, and Zoltners (1992) and Venkatesan and Kumar (2004). However, the total margin contributed may or may not have a similar relationship with marketing effort. Thus, we wish

[4] We subsequently test this assumption under the robustness checks section in Appendix C. It could be argued that a low income customer might order low margin items more frequently. However, it is also possible that a high income customer might order low margin items less frequently. Therefore, in the overall population, the relationship between purchase frequency and contribution margin per item will depend on the mix of high and low income consumers and their ordering behaviors and is thus an empirical issue.

to identify the response parameters associated with marketing efforts. Given the response parameters, marketing costs, and the total margin of a customer-channel segment, we can determine the level of marketing efforts that would maximize the profits from that segment. We do not impose an artificial marketing budget constraint because the one of the goals is to determine the optimal level of marketing efforts.

In the model, we include three different types of marketing instruments, namely, marketing mailer, promotional discount, and price. These instruments may have different elasticities. From a resource allocation perspective, however, we optimize only the number of mailers sent to a customer-channel segment. Such an assumption is consistent with industry practice. For example, J.C. Penney, which spends about $357 million on marketing mailers, makes such allocation decisions on the customer segments to whom mailers should be sent, based on the channels through which they shop (*J.C. Penney 10K Statement* 2006). It is also in line with the data provider's managerial practice.

We also assume the prices of the products and the discounts offered to customers as exogenous to the modeling system. Note that the aim of the study is to optimize resource allocation decision at the customer-channel segment level, so we do not pursue resource allocation at an individual customer level within each customer-channel segment.

We approach the resource allocation model in four steps. First, we estimate each of the four models (purchase frequency, purchase quantity, product return propensity, and contribution margin) for every customer-channel segment to get the response and shape parameters. We anticipate that the response parameters for a given marketing instrument will be different across different customer-channel segments. Second, using these response and shape parameters, we predict the different components of profit in the prediction window. Third, we evaluate the predictive ability of the model for each customer-channel segment. Finally, through simulation, using these estimated response and shape parameters and the cost of each type of marketing instrument for each customer-channel segment, we get the optimal values of $\sum_{i=1}^{n_k} m_{ik}$, the levels for marketing efforts that should be expended toward the customer-channel segment k. We compare the predicted profits derived from our model with those actually generated by the firm in a holdout sample. The difference reflects the profit improvement generated by our modeling approach.

Purchase Frequency Model

The two most popular methods to estimate purchase frequency are the Pareto/negative binomial distribution (NBD) (Schmittlein, Morrison, and Colombo 1987) and the generalized gamma model (Allenby, Leone, and Jen 1999). While the Pareto/NBD model assumes a Poisson distribution of customer transaction rate, the generalized-gamma model assumes a gamma distribution of customer transaction rate. The Pareto/NBD model is theoretically appealing, but its implementation requires tedious evaluation of multiple Gaussian hypergeometric functions. The availability of faster computing has partially resolved this problem. Some studies in marketing have successfully implemented the model (e.g., Fader, Hardie, and Lee 2005b; Reinartz and Kumar 2000). Fader, Hardie, and Lee (2005a) develop a derivative of the Pareto/NBD model, namely, the Beta Geometric (BG)/NBD model, which is significantly easier to implement, yet performs similar to the Pareto/NBD model. None of the studies that use the Pareto/ NBD and the BG/ NBD model in marketing, however, include the effects of covariates on purchase frequency. In addition to minimizing the bias in the parameter estimates, the inclusion of marketing covariates enables us to seek the optimal levels of such marketing covariates. The other derivatives of the Poisson class of models are the Conditional NBD and the Hierarchical Bayesian (HB) version of NBD (Jen, Chou, and Allenby 2003). While the Pareto/NBD and the BG/NBD are four-parameter models, the Conditional NBD and the HB NBD are two-parameter models, so their model fits are not comparable with those of the four-parameter models.

We use the BG/NBD model to estimate and predict customer purchase frequency, consistent with Fader, Hardie, and Lee (2005a). The BG/NBD model by Fader, Hardie, and Lee (2005a) has two distinct parts. The probability of a customer remaining active is captured by the BG part of the model. The transaction rate of a customer who is active is captured using the NBD part of the model. The probability of a customer remaining active and the transaction rate are assumed to be independent across customers.

We introduce three marketing covariates in the model: number of marketing mailers, the discount offered per item, and the average price of each item. The model derivation and estimation details are shown in Appendix B.

Purchase Quantity Model

While many studies investigate purchase frequency and purchase quantity separately, no study models the effect of marketing efforts on customer order size. Lewis (2006) and Lewis, Singh, and Fay (2006) model consumer order size in an online context as a function of shipping fees and find that the shipping fee structure can act as a motivation for consumers to increase their order sizes. The literature on customer basket size in the grocery industry has investigated the drivers of variation in customer basket size. The availability of a surprise coupon (e.g. in-store or shelf coupon) for a pre-planned purchase product category helps increase the size of a customer's basket (Heilman, Nakamoto, and Rao 2002).

To measure the average size of an order on a given purchase occasion, we apply the commonly used count data regression approach. Purchase quantity per order follows a Poisson distribution. However, the restrictive assumption of the mean and the variance being equal in the Poisson distribution cannot capture overdispersed data. The NBD distribution is an ideal substitute for the Poisson distribution when data exhibit overdispersion (Cameron and Trivedi 1998). We use the Conditional NBD model developed by Morrison and Schmittlein (1988) to model customer purchase quantity per order. We introduce three marketing covariates in the model: number of marketing mailers, the discount offered per item, and the average price of each item. The model derivation and estimation details appear in Appendix B.

Product Return Propensity Model

By modeling customer product return per order, we correct for the bias created in total margin derived from a customer when product returns are ignored. Prior studies in marketing do not account for the product return propensities of consumers and hence tend to overestimate purchase quantity, contribution margin, and thereby, the valuation of consumers. Product return per order is conceptually very similar to purchase quantity per order. However, we treat product return per order as a purely stochastic process. The model derivation and estimation details appear in Appendix B.

Contribution Margin Model

Contribution margin per item can be modeled at a customer level, order level and an item level. Prior studies have focused on the first two levels. Customer level approaches have used hierarchical models to capture the effect of firm-specific marketing intervention and customer-specific shopping traits. Venkatesan and Kumar (2004) use a panel data regression with lagged contribution margin to correct for model misspecification to capture the contribution margin generated by a customer at the order level. Fader, Hardie, and Lee (2005b) use a "regression to the mean" approach to capture the monetary value at the order level. Their modeling approach is superior to previous models because it incorporates heterogeneity across multiple orders for a customer and heterogeneity in average monetary value across customers. However, they model contribution margin at the order level and do not incorporate covariates in their model. We extend their model to incorporate the effect of marketing efforts and model the contribution margin per item.

We extend the Gamma-Gamma model used by Fader, Hardie, and Lee (2005b) to capture the effect of marketing efforts on average contribution margin per item for a given customer. We introduce two marketing covariates into the model: the number of marketing mailers and discount offered per item. We include the square of marketing mailers term to capture the decreasing return to scale. Although we include price in the purchase frequency and purchase quantity per order model as a covariate, we do not introduce it as a covariate in the contribution margin per item model. If a firm does not engage in dynamic pricing, the contribution margin and the price of a product are linearly related. Therefore, the introduction of price as a covariate in the model would lead to a nearly perfect prediction of the contribution margin by price. The model derivation and estimation are provided in Appendix B.

In the estimation window of duration 'T,' for a given customer with purchase frequency 'x', purchase quantity per order 'f', items returned per order 'h', and contribution margin per item 'w_{xf}', the total contribution margin derived from that customer will be $(f – h) * x * w_{xf}$. For the same customer in the prediction window of duration 't' the purchase frequency, purchase quantity per order, items returned per order, and expected contribution margin are given by 'Y', 'G', 'I', and 'U' respectively. The total margin derived from this customer in the prediction window is: (G – I) * Y * U. These values in the prediction window are obtained from the customer's responses to the firm's marketing efforts in the prediction window and the shape parameters of each model as estimated in the estimation window.

Data

We estimate the models using customer transaction data obtained from a large shoe and accessories manufacturing and marketing firm. The firm operates in the high-end market of the product categories and has been in this business for almost a century. The firm markets its products through physical stores, the Web, and catalogs The transaction data used in this study are from customers to whom the firm sells through its own retail network. The customer response file begins on January 1, 2003 and ends on August 7, 2005. The firm operates two different types of direct marketing campaigns. In addition to mailing 10 catalogs a year, the firm also offers promotional discounts semi-annually. The promotional flyers for these promotional discount campaigns are mailed to customers and prospects before the beginning of the sales period.[5] The prices of products in the promotion period are marked down by a fixed percentage. This information on the firm's marketing efforts is available at the individual customer level from January 1, 2004 until August 7, 2005. The sales promotion campaigns are implemented by the firm uniformly across all three channels. The prices of products are consistent across the channels during both the promotion period and the non-promotion period. In addition to the two types of direct marketing campaigns, the firm annually purchases about 20 to 25 inserts in national newspapers and magazines.[6] The firm advertises neither on television nor in the electronic media. A customer is tracked in the database by a unique customer identification key assigned to her when she makes her first

[5] Because marketing mailers are typically sent during holidays or seasonal time frames, seasonality and marketing mailers are typically highly correlated. Because marketing mailers are a key part of our model, we do not include seasonality in our model.

[6] Data on impressions generated through print media are available at the aggregate level and cannot be created at the individual level for use in individual level response models. These print media marketing efforts will influence customer behavior, but it is reasonable to assume that these effects are uniformly distributed across segments.

purchase with the firm. The database tracks customers' choices of the channel, the SKU purchased, the price paid, the date of transaction, and the date of return for every transaction. Each SKU bought by a customer constitutes one record in the customer response file. The product level file contains information on the product category, the SKU, the retail price, and the cost. The customer response file, marketing information file, and the product category file together constitute the dataset used in this study.

Of the 135 weeks of data, only 84 weeks beginning January 1, 2004 contain information about direct marketing efforts. We use the first 56 weeks of data as the estimation sample and the following 28 weeks of data as the holdout sample for testing the predictive validity of the models. We identify the cohort of first time buyers in the estimation window and use them for the estimation and prediction samples. There are 212,187 new customers with over half a million records in the estimation window. We calculate the discount offered to a customer for a given item as the retail price minus the dollar amount paid for that item by the customer. Similarly, we compute margin as the dollar amount paid for an item less the cost of the item. The price paid for a given SKU by customers will differ in the promotion and the non-promotion periods, thereby bringing variation in the discount variable in the data (see Equation A.4).[7]

The summary statistics on some of the key variables of the data are shown in Table 1. A cursory look at the table suggests that an average multichannel customer outspends an average single channel customer by a factor of at least two. A similar pattern is evident for purchase frequency and total number of items bought. Average store orders are larger than average Web and catalog orders. While customers buying in a store may have larger order sizes, customers shopping on the Web and through the catalog purchase the items with a larger contribution margin per item. The 'Web only' and 'catalog only' customers have smaller order sizes, but more than make up for this drawback by ordering more profitable items from the firm's product mix. The number of items returned per order is high for customers purchasing through the direct channels. This trend in return per order is in line with that reported by direct marketers. The firm's marketing efforts are higher for the 'multichannel' and 'catalog only' segments

[7] Data on shipping costs, which are relevant for the catalog and Web channels, and which vary by customer location, are not available. Because they are customer-transaction specific and are not a decision variable under the firm's control, their omission from the empirical analysis is not a serious issue.

than they for the ‘store only’ and ‘Web only’ segments. To correctly analyze the effects of marketing efforts, we only include marketing efforts expended before the last purchase of the customer in the data. ‘Multichannel’ and ‘store only’ customers exhibit the greatest tendency to shop across multiple categories.

< Insert Table 1 here >

Model Estimation, Simulation, and Results

Model Estimation and Simulation

We estimate the models using the maximum likelihood estimation method. This flexible approach allows the specification of user defined likelihood function, places constraints on the lower bounds of the shape and scale parameters, and selects appropriate starting values. We use an improvement in the log likelihood function by 1.0e-03 for at least two consecutive steps as the convergence criteria for the models. We gave different starting values to the model parameters and checked if the algorithm converged to same values of parameters for the log likelihood function. For obtaining the standard errors of the parameter estimates, we derived the Hessian matrix of each model. The Hessian matrix is the second derivative of the log likelihood function with respect to each parameter. We used the inverse of Hessian to calculate the information matrix. The square roots of the vector of the diagonal elements of this information matrix are the standard errors associated with the parameter estimates.

We also estimated the 2F1 Gauss hyper geometric function in the purchase frequency model. We used Equation (B.6) to evaluate the terms of hypergeometric series and Equation (B.5) to compute the value of the function for each customer. We evaluated the first 500 terms of the series. Although the computer on which the function was evaluated had a machine epsilon of 1.0e-300, the terms converged to zero much before the first 500 terms, so we used only the first 500 terms in the calculation of the hypergeometric function.

The practical estimation of each model can be performed by a manager on a statistical software package such as SAS or MATLAB. It could also be done in Excel if the number of observations is limited. The simulations can also be done in the same statistical software package or Excel.

Purchase Frequency Model Results

The results of purchase frequency model are presented in Table 2. The effect of marketing mailers on customer purchase frequency is positive and significant for each customer-channel segment. The results indicate that the ‘multichannel’ (2.314, $p < 0.001$) and the ‘store only’ (1.540, $p < 0.001$) customers are most responsive to marketing mailers. The effect of marketing mailers on purchase frequency of the ‘catalog only’ (1.134, $p < 0.001$) and the ‘Web only’ (1.432, $p < 0.001$) customers is positive and significant. The effect of an additional mailer on purchase frequency of customers increases at a decreasing rate. The average response to an additional mailer is a 2.5% increase in purchase frequency for up to seven mailers. Customers’ response to marketing mailers flattens beyond this level.

< Insert Table 2 >

The effect of the prices of SKUs bought by customers on their purchase frequency is in the expected direction. The anticipated spending on planned purchases by a consumer is endogenous to her disposable income. Thus, the relationship between the price of SKUs bought (consumer spending) and purchase frequency is likely to be negative. The ‘store only’ customers exhibit the highest degree of spending constraint on each shopping occasion (-0.130, $p < 0.001$). The effect of price on the purchase frequency of ‘Web only’ (-0.099, $p < 0.001$) and ‘catalog only’ (-0.073, $p < 0.001$) customers is also negative and significant. Previous studies show that customers who order through the direct channels tend to spend less on each purchase occasion to mitigate the risk of using a direct channel. ‘Multichannel’ customers, however, do not exhibit a similar spending constraint (-0.014, $p > 0.05$).

The results also suggest that discounts to ‘store only’ (0.363, $p < 0.001$), ‘multichannel’ (0.319, $p < 0.001$), and ‘Web only’ (0.124, $p < 0.001$) customers are positively associated with purchase frequency. The results do not indicate a significant effect of discount on purchase frequency for the ’catalog only’ segment ($p > 0.05$).

Purchase Quantity Model Results

The results of the purchase quantity model are reported in Table 3. They indicate that marketing mailers positively influence the purchase quantity per order for each customer segment. The effect of marketing mailers on purchase quantity per order is high for customers of the 'multichannel' (1.513, $p < 0.001$), the 'catalog only' (1.441, $p < 0.001$) and the 'Web only' (1.403, $p < 0.001$) segments. These findings indicate that in response to a firm's marketing efforts, multichannel customers not only purchase more frequently, but also order more items per purchase occasion than do single channel customers. The effect of marketing mailers on purchase quantity per order of 'store only' customers is also positive significant (1.098, $p < 0.001$), although it is the lowest among all the segments.

< Insert Table 3 here >

Price has a negative and significant relationship with purchase quantity per order for 'store only', 'catalog only' and 'Web only' customers. This result indicates that when customers order more expensive items, they purchase fewer items on that purchase occasion. This relationship is strongest for the 'Web only' customer segment (-0.109, $p < 0.001$). Evaluating this result together with the results of the purchase frequency model suggests that when purchasing more expensive items, 'catalog only', 'store only', and 'Web only' customers not only purchase less frequently in a given time window, but also have smaller order sizes on each purchase occasion. The relationship between price and purchase quantity for 'multichannel' customers is negative, but statistically insignificant (-0.003, $p > 0.05$). This finding may be due to the relatively smaller sample size for the segment. However, a potential explanation is that 'multichannel' customers have significantly higher income than 'single channel' customers (Kushwaha and Shankar 2007), so price does not have a significant effect on quantity for them.

The effect of discount on purchase quantity per order is similar to its effect on purchase frequency. The 'store only' and 'multichannel' customer-channel segments are highly responsive to discounts ($p < 0.001$). Other customer-channel segments also have positive and significant response coefficients to discounts ($p < 0.001$), but these coefficients are smaller than those for the 'store only' and 'multichannel' segments.

Product Return Propensity Model Results

The shape and scale parameters of the Conditional NBD model for product return propensity per order are presented in Table 4. All the parameter estimates are significant ($p < 0.01$) and the model fits are good. Therefore, the predicted values help accurately estimate the net purchases in the prediction window.

< Insert Table 4 here >

Contribution Margin Model Results

The results of the contribution margin model suggests an inverted 'U' relationship between the number of marketing mailers received and the contribution margin of an average item from each customer-channel segment. 'Store only' (49.929, $p < 0.001$) and 'multichannel' (93.104, $p < 0.001$) customers, who receive marketing mailers are more likely to purchase items with higher contribution margins than those who do not receive the mailers. However, the large and negative coefficient of the square of marketing mailers term for 'store only' (-2.642, $p < 0.001$) and 'multichannel' (-4.329, $p < 0.001$) customers also indicate that this increasing effect of marketing mailers on contribution margin decreases at a significantly faster rate than it does for other customers. The extremely large effect of marketing mailers on the contribution margin per item of 'store only' and 'multichannel' (for store specific purchases) customers could be partly attributed to store-specific unobserved effects. Store-specific unobservables such as personal selling efforts and customer interactions of the sales staff could be providing an extra impetus for the 'store only' customer to make purchases of high-margin items. The relationship between marketing mailers and contribution margin per item is similar for customers using direct channels, although it is not as strong as it is in the case of 'store only' or 'multichannel' customers. The results of the model are presented in Table 5.

< Insert Table 5 here >

The results also suggest a strong positive relationship between the amount of discount offered and the contribution margin per item for the 'store only' (14.525, $p < 0.001$), 'catalog only' (4.868, $p < 0.001$), and 'Web only' (2.815, $p < 0.05$) customer-channel segments.

Optimization Model Results

Customers' channel preferences and choices may be dynamic and evolve over a period of time (Ansari, Mela, and Neslin 2008). All multichannel customers begin as 'single channel' customers and over a period of time, may adopt additional channels to transition to the 'multichannel' state (Venkatesan, Kumar, and Ravishanker 2007). In our data, some customers who are classified as single channel customers in the estimation time window may transition to the 'multichannel' segment. In the prediction window, we determine their probabilities of transition to multichannel customers. The numbers of customers who transitioned from one segment to another from the estimation to the prediction window are presented in Table 6. Only 2%, 1%, and 4% of 'catalog only,' 'store only' and 'Web only' customers transitioned into the 'multichannel' state. Even so, to account for the dynamic nature of these segments, we incorporate the segment transition probabilities of customers in our optimization model. We adjust the estimates of each segment profit by the retention probability of customers in that segment.

< Insert Table 6 here >

The results of the optimization model are presented in Table 7. The results suggest that if the allocated budget for this cohort of customers is increased by 157%, the profits from the entire cohort of customers can be improved by as much as 33%. The optimization results suggest that the budgets for 'store only' segment and 'multichannel' segment should be increased by 171% and 180%, respectively. The results of the optimization model are representative of the responsiveness of these segments to marketing mailers, especially with respect to their purchase frequency, purchase quantity, and contribution margin. The results in Tables 2, 3, and 5 suggest that the 'multichannel' and 'store only' customer-channel segments are most responsive to the firm's marketing mailers. Before optimal allocation, the 'store only' and 'multichannel' segments were being allocated 78% and 4% of the total budget. After optimal allocation, these segments now receive 82% and 5% of the substantially increased budget.

< Insert Table 7 here >

The optimal allocation results also suggest that the budgets for the 'catalog only' segment and the 'Web only' segment should be increased by 67% and 128%, respectively. The suggested allocation for these segments reflects their responsiveness to the firm's marketing mailers as revealed by the results presented in Tables 2, 3, and 5. Before optimization, the 'catalog only' and 'Web only' segments were each allocated 9% of the marketing budget. After optimization, the results suggest that these segments should be allocated only 6% and 8%, respectively, of the substantially increased budget.

The optimization results presented in Table 7 suggest that in response to the recommended increase in marketing budget, the profit from this cohort of customers over the 28-week prediction window would increase from $2.94 million to $3.86 million. This increase in profit reflects the extraction of an additional $5 from each customer in this cohort. This increase, when extrapolated over a 52-week period for the entire customer base, translates to additional profit of over $6.56 million. The share of profit (as a percentage of total profits) contributed by 'Web only' and 'multichannel' customers increases from 8% and 5% to 10% and 9%, respectively. This increase reflects a 68% and 138% improvement in the absolute dollar value of profits from the 'Web only' and 'multichannel' segments, respectively. Although the absolute dollar value of profits from the 'store only' and 'catalog only' segments increase by 24% and 21%, respectively, the relative (percentage of total profits) contributions from these segments to the firm's profit declines to 75% and 6%, respectively. The results also indicate that for every one dollar increase in the marketing budget, the additional profits contributed by the 'catalog only,' 'store only,' 'Web only' and 'multichannel' segments are $2.55, $1.47, $4.79 and $9.75, respectively. For the entire cohort of these customers, this reallocation represents an increase of $2 in profit for every additional dollar invested in marketing mailers. The proposed increase and reallocation of marketing efforts would drive customers to adopt additional channel(s), increasing the size of the multichannel segment and improving overall profitability. Because we do not model customers' channel preferences directly as a function of marketing efforts, our results are conservative and the potential profit improvement may be greater than that suggested by our models.

The shape of the profit function against marketing mailers for each customer-channel segment is shown in Figure 2. The optimal profit per customer for the entire customer cohort is $18.45, which is achieved when a customer receives 6.37 mailers. Beyond this level of marketing efforts, the profit starts to decline because the increase in marketing cost outweighs the customer's responsiveness to those marketing efforts. The profit functions for 'store only' and 'catalog only' customers are similar in shape, decreasing almost symmetrically after about seven mailers. However, the optimal level of profit per customer (number of marketing mailers per customer) is somewhat lower for the ‘store only’ segment ($16.31 [6.19]) than it is for the ‘catalog only’ segment $18.43 ([5.88]). The shapes of the profit functions for the ‘Web only’ and ‘Multichannel’ segments are somewhat similar, but the optimal profits and the optimal number of marketing mailers per customer are quite different. The optimal profit (optimal number of marketing mailers) per customer is much higher for the ‘multichannel’ segment ($57.84 [9.72]) than it is for the ‘Web only’ segment ($27.93 [7.66]). Because the ‘store only’ segment is the largest segment, the optimal profit and number of marketing mailers per customer for the entire customer cohort is closer to those for the ‘store only’ segment. Nevertheless, the other segments, particularly, the ‘multichannel’ segment significantly boost the profit per customer of the cohort.

< Insert Figure 2 here >

Our results are robust to model assumptions, alternate operationalizations of variables, and alternate potential explanations of the phenomenon under investigation. The details of robustness checks appear in Appendix C.

Model Validation

Predictive Validation: ‘Catalog Only’ Segment

For ‘catalog only’ customers, charts of the predicted and actual values of the profit components are presented in Figure 3. For the purchase frequency model, the beta geometric part and the negative binomial part of the model fit the data very well. The error in predicting the dropout rate is only 2%. The NBD part of the model also fits well for customers making one purchase. The predicted value of customers making only two repeat purchases underestimates the actual figure by 1.5%. The model fits for the purchase quantity model are similar to those for the purchase frequency model. The model performs very well for customers with smaller order sizes, but the errors swell to about 1% for larger order sizes. In the model capturing product returns per order, customers who are likely to return at least one item are under-represented by almost 4%. Toward the right tail of the distribution, the predicted values are slightly overestimated.

< Insert Figure 3 here >

Predictive Validation: 'Store Only' Segment

The model fits of the 'store only' customer-channel segment are extremely good as shown in Figure 4. The purchase frequency model performs very well in predicting the customer dropout rate. The NBD part of the model also captures the data very well and predicts the purchase frequency of customers within 2.5% of the actual values. The model fit of the purchase quantity per order model is also excellent with the prediction results within half a percent of the actual values across the distribution. The prediction results of the items returned per order model are also on the optimistic side. It under-represents the percentage of customers who will return only one item for every order they place by 3%. However, it also over represents customers returning more than one item by 4%.

< Insert Figure 4 here >

Predictive Validation: 'Web Only' Segment

The model fits of the 'Web only' segment, shown in Figure 5, are very similar to those for the 'catalog only' customer-channel segment. The inherent similarity in behavior across customers who use only direct channels is evident. We observe a maximum deviation of 2% between the actual and predicted values of purchase frequency for customers making only one repeat purchase. The product returns propensity model predicts fewer people who are likely to return an item for every order they place. The 'Web only' customer-channel segment purchases items with the highest contribution margin per item.

< Insert Figure 5 here >

Predictive Validation: 'Multichannel' Segment

The results for the multichannel customer-channel segment appear in Figure 6. The BG/NBD model used for estimating purchase frequency assumes that customer drop out will occur at some point. However, multichannel customers are those customers who have ordered at least twice using at least two different channels. Because every customer in the multichannel customer segment ordered more than once, the BG/NBD model does not capture the phenomenon as well. Similarly, the purchase quantity per order model which measures items purchased per order only in the repeat orders, does not adequately capture the phenomenon. To circumvent this problem, we rescaled the measure of order size by subtracting one to meet the assumptions of BG/NBD model. This rescaling resulted in a significantly better fit for the purchase frequency model and moderate improvement in fit for the purchase quantity model. The fit of the purchase quantity model remains moderate with highest error in prediction--up to 10% of actual value. The fit of the product return propensity model is very good with the predicted values remaining within 4% of the actual values.

< Insert Figure 6 here >

Predictive Validation: Contribution Margin Model

Because contribution margin is a point estimate, the predictive validity of the contribution margin model for all the segments is shown in a single figure, namely, Figure 7. The predicted and actual values are close to each other for all the segments. For the ‘catalog only’ segment, the predicted margin per item is $78.62 and is within 2% of the actual value of $79.75. For the ‘store only’ segment, predicted contribution margin per item is $55.23 compared to the actual value of $61.35. This underestimation is likely because of large variance in margin per item for ‘store only’ customers. This finding is evident from the fact that the shape parameter (q) of the gamma distribution, which captures heterogeneity across customers for this segment, is one of the largest among all customer-channel segments. For the ‘Web only’ segment, the predicted margin is $78.58 and compares well with the actual value of $80.92. For the ‘multichannel’ segment, the predicted margin per item is $65.78 compared to the actual value of $68.03.

< Insert Figure 7 here >

Implications, Limitations, Future Research and Conclusion

Managerial Implications

The proposed model has important managerial implications. They show that an ‘a priori’ segmentation of customers based on their channel choices is theoretically and managerially relevant. The model can help managers identify how much marketing efforts should be expended for each channel segment. By decomposing profits from the customer into multiple components of purchase, the model enables managers to identify the influence of marketing efforts on the customer's purchase frequency, purchase quantity, and contribution margin. An attractive feature of the model is that it is based on predicted future profits. The model can serve as a decision support tool for marketing resource allocation decisions as well as for the design of marketing communication. The model is generalizable and can be implemented in variety of contexts.

To show the usefulness of the decomposition approach of the model and draw managerial insights into the differential responses of the customer-channel segments by the profit components, we performed a post-hoc analysis. Of the three covariate-dependent components of profit, namely, purchase frequency, purchase quantity, and contribution margin, we held two variables at their actual values, while setting the third at the predicted value for calculating the improvement in profits. If AcPF, AcPQ, AcCM, and PrPF, PrPQ, PrCM are the actual and predicted purchase frequencies, purchase quantities, and contribution margins, respectively, the actual firm profit is given by: AcPF*AcPQ*AcCM. The difference in the profits realized by replacing the actual value of one component by its predicted counterpart is the dollar value of the contribution made by the component. Thus, the additional profit generated by increase in purchase frequency in response to greater number of mailers is given by (PrPF*AcPQ*AcCM - AcPF*AcPQ*AcCM). For the entire customer cohort, the results reveal that 54%, 18%, and 28% increases in profits are contributed by improvements in contribution margin, purchase quantity, and purchase frequency, respectively, which translate into per customer gains of \$2.50, \$0.79, and \$1.29, respectively. Thus, through our modeling approach, we are able to differentiate between improvements in purchase quantity and contribution margin. Previous modeling approaches could only suggest that improvement in the dollar value of an order would contribute an improvement in aggregate profits.

The average profits generated by a customer from each segment and the decomposition of those profits are shown in Table 8. The increase in average profits is highest for 'multichannel' customers. About 60% of this increased profit comes from purchases of high-margin items and the remaining 40% of this increase arises due to the purchase of more items and more frequent purchases. The decomposition for the ‘store only’ segment also reveals similar results. However, for ‘Web only’ customers, an increase in profits is primarily contributed by increases in purchase frequency and purchase quantity. For ‘catalog only’ customers, the hike in profits comes approximately equally from all the three components. Thus, with regard to the decomposition of additional profits to optimized marketing efforts, ‘store only’ and ‘multichannel’ customers are different from customers who primarily use direct channels. These findings can also help managers tailor the content of their marketing communication to each segment. For example, the marketing mailers targeted at ‘multichannel’ and ‘store only’ segments could feature more high-contribution items. Similarly, the marketing mailers targeted at customers using only the direct channels could feature more products and could be sent more frequently.

< Insert Table 8 here >

We present a strategic summary of the key findings in Table 9 and discuss the key insights. The ‘multichannel’ and ‘Web only’ segments produce very high returns on investment. Every additional marketing dollar yields profits of \$10 and \$5 from the ‘multichannel’ segment and the ‘Web only’ segment, respectively. These segments exhibit significant increases in both absolute and relative profits due to reallocation. However, the high financial returns from both these segments stem from very different customer behaviors. The high return of investment from multichannel customers is attributable to the high responsiveness of multichannel customers to marketing mailers, particularly, with respect to their purchase frequency and *contribution margin*. For this segment, the high effect of marketing mailers on purchase quantity is mitigated by its high propensity to return merchandise. However, for the ‘Web only’ segment, the significant profit improvement primarily arises from its high responsiveness to marketing mailers with respect to purchase frequency and *purchase quantity*. For this segment, the significant effect of mailers on purchase quantity is augmented by its low tendency to return products.

< Insert Table 9 here >

The ‘catalog only’ and ‘store only’ segments produce low-moderate returns on investment. Every additional marketing dollar yields profits of \$3 and \$1 from the ‘catalog only’ segment and the ‘store only’ segment, respectively. Although these segments have lower relative contributions to optimal profits than to actual profits, their relative importance is magnified by their sizes. Despite the similar response profiles, the behaviors of customers in these segments are different. The purchase behavior of 'store only' customers is similar to that of 'multichannel' customers, exhibiting high responsiveness to marketing mailers with respect to their contribution margin. At the same time, 'catalog only' customers are similar to 'Web only' customers in the relative importance of each profit component to their response profiles.

These managerial insights suggest that customers of non-traditional ('multichannel' and 'Web only') channels offer high potential for profit improvement. These findings may be partly explained by the demographic profiles of customers constituting these segments (Kushwaha and Shankar 2007). Another possible explanation is the manner in which customers in these segments weigh and process information. We also show that the behavior of customers using direct channels ('catalog only' and 'Web only') can be triggered in way different from those customers who prefer to use bricks-and-mortar stores. These differences in profit potential and purchase behavior across the customer-channel segments can be used to design effective marketing communication strategies.

Limitations and Future Research

The research has certain limitations that could be addressed by future research. First, although we test for the endogeneity of customer channel choice and it is not a problem in our data based on our test, channel choice could be a function of marketing effort in other contexts. Models developed for such contexts could capture this possibility. Second, we treated segments of all combinations of channel choices to be one multichannel segment. Future research could examine if there are any differences in response to marketing efforts across different types of multichannel segments (e.g., store and the Web, and store and catalog). Third, we focused on the number of marketing mailers as the key decision variable because it was the most important actionable variable to the company that provided the data. Future research could extend the decision variables to other variables such as price and discount. Finally, firms could use this model as the basis for real world experiments of marketing resource allocation decisions.

Conclusion

In conclusion, we addressed three important research questions and proposed a model for the optimal allocation of marketing efforts across multiple customer-channel segments. We developed and estimated a set of marketing covariate-driven stochastic models for purchase frequency, purchase quantity, and contribution margin for each customer-channel segment. Based on the parameter estimates from the models, we then derived the optimal marketing effort allocation to each customer-channel segment for a future holdout period, using simulation. The optimization model can be implemented using Excel software. The results from the application of our model show that firms can substantially improve their profits by optimally allocating their marketing resources to these customer-channel segments based on the heterogeneous responses of these segments to the firm's marketing efforts across the different components of profit. The results suggest that the multichannel segment offers greater profit potential than do other segments primarily through high contribution margin, allowing greater allocation of marketing efforts; however, overall allocation is primarily driven by the relative sizes of the customer-channel segments.

References

Achabal, D.D., M. Badgett, J. Chu, and K. Kalyanam (2005), "Cross-Channel Optimization: A Strategic Roadmap for Multichannel Retailers," *White Paper*, IBM Consulting Services

Allenby, Greg, Robert Leone, and Lichung Jen (1999), "A Dynamic Model of Purchase Timing with Application to Direct Marketing," *Journal of the American Statistical Association*, 94(446), 365-374.

Ansari, Asim, Carl Mela, and Scott Neslin (2008), "Customer-Channel Migration," *Journal of Marketing Research*, forthcoming.

Assael, Henry. 1998. *Consumer Behavior and Marketing Actions.* South-Western College Publication, Cincinnati, OH.

Balasubramanian, S., R. Raghunathan, and V. Mahajan (2005), "Consumers in a Multi-channel Environment: Product Utility, Process Utility, and Channel Choice," *Journal of Interactive Marketing*, 19(2), 12-30.

Bart, Yakov, Venkatesh Shankar, Fareena Sultan, Glen L. Urban (2005), "Are the Drivers and Role of Online Trust the Same for all Web Sites and Consumers? A Large-Scale Exploratory Empirical Study," *Journal of Marketing,* 69(4), 133-152.

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

Cameron, A. Colin and Pravin K. Trivedi (1998), *Regression Analysis of Count Data*, Cambridge University Press: Cambridge, UK.

Doubleclick. 2004. *2004 Holiday Season Multi-Channel Retail Research.* Doubleclick Inc, NewYork, NY.

Elsner, Ralf, Manfred Kraft, and Arnd Huchzermeier (2003), **"**Optimizing Rhenania's Direct Marketing Business Through Dynamic Multilevel Modeling (DMLM) in a Multicatalog-Brand Environment," *Marketing Science*, 23(2), 192-106.

Fader, Peter S., Bruce G. S. Hardie, and Ka Lok Lee (2005a), "Counting Your Customers the Easy Way: An Alternative to Pareto/NBD Model," *Marketing Science*, 24(2), 275-284.

_____, _____, and _____ (2005b), "RFM and CLV Using Iso-Value Curves for Customer Base Analysis," *Journal of Marketing Research*, 42(4), 415-430.

_____, _____, and _____ (2005c), "Implementing the BG/NBD Model for Customer Base Analysis in Excel," http://www.brucehardie.com/notes/004/, Accessed on May 15, 2006

Fenvessy, Stanley. 1992. Fulfillment planning: an overview. in *The Direct Marketing Handbook.* Edward Nash. ed. McGraw-Hill, New York.

Heilman, Carrie M, Kent Nakamoto, and Ambar Rao (2002), "Pleasant Surprises: Consumer Response to Unexpected Instore Promotions," *Journal of Marketing Research*, 39(2), 242-251.

Hess, James D. and Glenn E. Mayhew, (1997), “Modeling Merchandise Returns in Direct Marketing,” *Journal of Direct Marketing,* 11(2) 20-35.

Inman, Jeffery, Venkatesh Shankar, and Rosellina Ferraro (2004), “The Roles of Channel-Category Associations and Geodemographics in Channel Patronage Decision,” *Journal of Marketing*, 68(2), 51-71.

J.C. Penney 10K Statement. 2006. J.C. Penney Company Inc: Form 10K. http://thomson.mobular.net/thomson/7/2348/2579/. Accessed on May 26, 2007.

Jen, Lichung, Chien-Heng Chou, and Greg M. Allenby (2003), A Bayesian Approach to Modeling Purchase Frequency,” *Marketing Letters,* 14(1) 5-20. Kumar V. and Rajkumar Venkatesan (2005), “Who are Multichannel Shoppers and How do they Perform?: Correlates of Multichannel Shopping Behavior,” *Journal of Interactive Marketing*, 19(2), 44-61.

Kushwaha, Tarun and Venkatesh Shankar (2007), “Single Channel vs. Multichannel Customers: Determinants and Value to Retailers,” Working Paper, Texas A&M University, College Station, TX 77843.

Lewis, Michael (2006), “Effect of Shipping Fees on Purchase Quantity, Customer Acquisition and Store Traffic,” *Journal of Retailing*,82(1), 13-23.

______, Vishal Singh, and Scott Fay (2006), “An Empirical Study of Impact of Non Linear Shipping Fees on Purchase Incidence and Expenditure Decisions,” *Marketing Science*, 25(1), 51-64.

Libai, B., D. Narayandas, and C. Humby, (2002), "Toward an Individual Customer Profitability Model: A Segment-Based Approach," *Journal of Service Research*, 5(1), 69-76.

Mantrala, Murali K., Prabhakant Sinha and Andrew Zoltners (1992), “Impact of Resource Allocation Rules on Marketing Investment-Level Decisions and Profitability,” *Journal of Marketing Research*, 29(2), 162-175.

Messinger, Paul and Chakravarthi Narasimhan (1997), “A Model for Retail Format Based on Consumers’ Economizing of Shopping Time,” *Marketing Science,* 16(1) 1-23.

Morgan, Robert M. and Shelby D. Hunt (1994), “The Commitment-Trust Theory of Relationship Marketing,” *Journal of Marketing,* 58(3) 20-38.

Morrison, D.G. and Schmittlein, D.C., (1988), "Generalizing the NBD Model for Customer Purchases: What Are the Implications and is it Worth the Effort?," *Journal of Business and Economic Statistics*, 6(2), 145-166.

Myers, Joseph B., Andrew D. Pickersgill, and Evan S. VanMetre (2004), “Steering Customers to the Right Channels,” *McKinsey Quarterly*, 4, 36-48.

Naik, Prasad A, Murali Mantrala K. and Allen G. Sawyer (1998), “Planning Media Schedules in the Presence of Dynamic Advertising Quality,” *Marketing Science*, 17(3), 214-235.

Neslin, S.A., D. Grewal, R. Leghorn, V. Shankar, M. L. Teerling, J. S. Thomas, P. C. Verhoef (2006), “Challenges and Opportunities in Multichannel Management,” *Journal of Service Research*, 9(2), 95-113.

_________ and Venkatesh Shankar (2007), “Key Issues in Multichannel Customer Management: Current Knowledge and Future Directions,” Working Paper, Dartmouth College, NH.

Rangaswamy, Arvind, Prabhakant Sinha, and Andris Zoltners (1990), “An Integrated Model Based Approach for Sales Force Structuring,” *Marketing Science*, 9(4) 279-298.

Reinartz, Werner and V. Kumar (2000), “On the Profitability of Long-Life Customers in a Noncontractual Setting: An Empirical Investigation and Implication for Marketing,” *Journal of Marketing*, 64(4) 17-35.

______, Jacquelyn S. Thomas, and V. Kumar (2005), "Balancing Acquisition and Retention Resources to maximize Customer Profitability," *Journal of Marketing,* 69(1), 63-79.

Schmittlein, David C., David G. Morrison, and Richard Colombo (1987), “Counting Your Customers: Who they Are and What Will they Do Next,” *Management Science*, 33(1), 1-24.

Shankar, Venkatesh and Marie Hollinger (2007), “Online and Mobile Advertising: Current Scenario, Emerging Trends, and Future Directions,” *Marketing Science Institute* Special Report, 07-206.

Thomas, Jacquelyn and Ursula Sullivan (2005), “Managing Marketing Communication with Multichannel Customers,” *Journal of Marketing*, 69(4), 239-251.

Venkatesan, Rajkumar and V. Kumar (2004), “A Customer Lifetime Value Framework for Customer Selection and Resource Allocation Strategy,” *Journal of Marketing*, 68(4), 106-126.

_______, _______, and N. Ravishanker (2007), “Multichannel Shopping: Causes and Consequences,” *Journal of Marketing*, 71(2) 114-132.

Verhoef, Peter C. and Bas Donkers (2005), “The Effect of Acquisition Channels on Customer Loyalty and Cross-Buying,” *Journal of Interactive Marketing*, 19(2), 31-43.

_______, S. A. Neslin, and B. Vroommen (2007), “Multichannel Customer Management: Understanding the Research Shopper Phenomenon,” *International Journal of Research in Marketing*, 24(2) 129-148.

Villaneueva, Julian, Shijin Yoo, and Dominique M. Hanssens (2008), “The Impact of Acquisition Channels on Customer Equity,” *Journal of Marketing Research*, forthcoming.

Table 1
Means and Standard Deviations of Key Variables in the Estimation Window of Data

	Catalog Only	Store Only	Web Only	Multichannel	All Customers
Cohort Size	13,783	180,305	13,955	4,144	212,187
Total Spending ($)	237.96[a]	237.37[a]	218.58	525.53	241.98
	(262.90)	(281.39)	(298.32)	(425.31)	(287.72)
Total Margin ($)	155.33	147.78	142.09	329.42	151.44
	(169.65)	(175.65)	(250.77)	(265.96)	(185.04)
Number of Orders	1.25	1.22	1.19	2.64	1.25
	(0.62)	(0.77)	(0.53)	(1.17)	(0.79)
Total Items	2.16	2.90	1.99	5.51	2.84
	(2.31)	(3.88)	(1.96)	(4.63)	(3.75)
Number of Categories Bought	1.20	1.49	1.18	1.79	1.46
	(0.45)	(0.61)	(0.43)	(0.71)	(0.60)
Order Size	1.73	2.37	1.67	2.09	2.27
	(1.46)	(1.92)	(1.36)	(1.33)	(1.86)
Return per Order	0.22	0.08	0.16	0.20	0.10
	(0.54)	(0.36)	(0.47)	(0.46)	(0.39)
Spending per Item ($)	120.84[b]	98.91	123.01[b]	108.11	102.10
	(77.24)	(64.61)	(214.72)	(58.76)	(83.25)
Margin per Item ($)	79.75[c]	61.46	80.93[c]	68.03	64.06
	(49.46)	(41.75)	(208.34)	(37.89)	(67.55)
Discount per Item ($)	2.80	3.84[d]	4.52	3.66[d]	3.81
	(10.82)	(9.83)	(17.61)	(7.27)	(10.55)
No. of Mailers	3.38	2.27	2.84	3.91	2.41
	(2.81)	(2.57)	(2.84)	(2.65)	(2.63)

Note: Standard deviations in parentheses. The means of all the variables except the pairs denoted by [a, b, c, and d] are statistically different across the customer-channel segments p < 0.05 or better).

Table 2
Results of Purchase Frequency Model

Variable/ Parameters	Catalog Only	Store Only	Web Only	Multichannel
r	0.563***	0.144***	0.220***	0.147***
	(0.006)	(0.002)	(0.003)	(0.003)
a	18.064***	20.165***	18.020***	12.338***
	(0.406)	(0.225)	(0.191)	(0.369)
b	1.826***	2.522***	9.355***	16.734***
	(0.161)	(0.190)	(0.249)	(0.250)
Intercept	3.049***	0.362***	0.987***	1.963***
	(0.021)	(0.041)	(0.029)	(0.053)
Mailers	1.134*	1.540***	1.432***	2.314***
	(0.449)	(0.320)	(0.139)	(0.657)
Price	-0.073***	-0.130***	-0.099***	-0.014
	(0.015)	(0.014)	(0.017)	(0.011)
Discount	0.105	0.363***	0.124***	0.319***
	(0.152)	(0.035)	(0.021)	(0.043)
Log-Likelihood	-19408.210	-223032.375	-15184.407	-11608.865

Note: Standard errors in parentheses. * p < 0.05, ** p < 0.01, *** p < 0.001.

Table 3
Results of Purchase Quantity Model

Variable/ Parameters	Catalog Only	Store Only	Web Only	Multichannel
c	0.514**	0.252***	0.857***	0.423***
	(0.159)	(0.033)	(0.125)	(0.051)
Intercept	1.962***	3.662***	2.007***	2.239***
	(0.023)	(0.024)	(0.026)	(0.016)
Mailers	1.441***	1.098***	1.403***	1.513***
	(0.192)	(0.217)	(0.117)	(0.195)
Price	-0.073***	-0.043***	-0.109***	-0.003
	(0.005)	(0.007)	(0.006)	(0.003)
Discount	0.122***	0.624***	0.263***	1.140***
	(0.017)	(0.023)	(0.020)	(0.016)
Log-Likelihood	-10110.221	-133470.432	-8330.783	-5176.754

Note: Standard errors in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table 4
Results of Product Return Propensity Model

Parameters	Catalog Only	Store Only	Web Only	Multichannel
d	0.298***	0.199***	0.373***	0.366***
	(0.027)	(0.015)	(0.016)	(0.013)
μ	26.082***	29.043***	24.216***	27.703***
	(0.113)	(0.142)	(0.095)	(0.045)
Log-Likelihood	-8530.994	-65405.296	-6837.797	-2769.307

Note: Standard errors in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table 5
Results of Contribution Margin Model

Variable/ Parameters	Catalog Only	Store Only	Web Only	Multichannel
p	2.831***	1.248***	2.941***	1.448***
	(0.015)	(0.006)	(0.016)	(0.017)
q	8.973***	11.347***	8.252***	13.787***
	(0.024)	(0.030)	(0.021)	(0.045)
Intercept	100.624***	232.476***	76.585***	151.064***
	(5.816)	(10.030)	(4.583)	(4.263)
Mailers	17.124***	49.929***	9.769***	93.104***
	(1.825)	(4.621)	(1.634)	(2.834)
Mailers Square	-0.761***	-2.642***	-0.310*	-4.329***
	(0.198)	(0.582)	(0.118)	(0.393)
Discount	4.868***	14.525***	2.815*	4.634
	(1.131)	(3.126)	(1.116)	(2.511)
Log-Likelihood	-72643.665	-932191.787	-73709.974	-20368.314

Note: Standard errors in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Table 6
Customer Transition across Segments

		Prediction Window				
		Catalog	Store	Web	Multichannel	Total
Estimation Window	Catalog only	13,485	0	0	298	13,783
	Store only	0	178,902	0	1,403	180,305
	Web only	0	0	13,418	537	13,955
	Multichannel	0	0	0	4,144	4,144
	Total	13,485	178,902	13,418	6,382	212,187

Note: The diagonal elements represent retention within a customer-channel segment, while the off-diagonal elements represent transition across segments.

Table 7
Optimization Results

Segment	**Budget**			**Profits**		
	$ Value	% of Total	% Change Over Actual	$ Value	% of Total	% Change Over Actual
Actual						
Catalog only	26,153	9.07		209,378	7.12	
Store only	225,098	78.08		2,352,986	80.01	
Web only	24,837	8.61		223,058	7.59	
Multichannel	12,210	4.24		155,313	5.28	
Total	288,297	100.00		2,940,736	100.00	
Suggested						
Catalog only	43,621	5.87	66.79	254,012	6.49	21.32
Store only	609,149	81.94	170.62	2,917,251	74.51	23.98
Web only	56,533	7.60	127.62	374,799	9.57	68.03
Multichannel	34,138	4.59	179.59	369,133	9.43	137.67
Total	743,440	100.00	157.87	3,864,119	100.00	33.14

Table 8
Profitability Decomposition

	Actual Profits ($)	**Optimal Profits ($)**	**Change ($)**	**Purchase Frequency (%)**	**Purchase Quantity (%)**	**Contribution Margin (%)**
Catalog only	15.19	18.43	3.31	34.19	35.25	30.55
Store only	13.15	16.31	3.15	26.64	14.15	59.21
Web only	16.62	27.93	11.31	44.90	41.81	13.29
Multichannel	24.34	57.84	33.50	23.50	16.44	60.06
All customers	13.84	18.45	4.59	28.19	17.31	54.49

Table 9
Strategic Summary of Findings

	Size of Segment	Optimal Budget Change	Optimal Profit Change	Return on Investment	Component of Profit		
					Purchase Frequency	Purchase Quantity	Contribution Margin
Catalog only	Moderate	Low	Moderate	Moderate	Moderate	Moderate	Moderate
Store only	Very Large	High	Moderate	Small	Moderate	Low	High
Web only	Moderate	Moderate	High	Large	High	High	Low
Multichannel	Small	Very High	Very High	Very Large	Moderate	Low	High

Figure 1
Conceptual Framework for Multichannel Resource Allocation Decisions

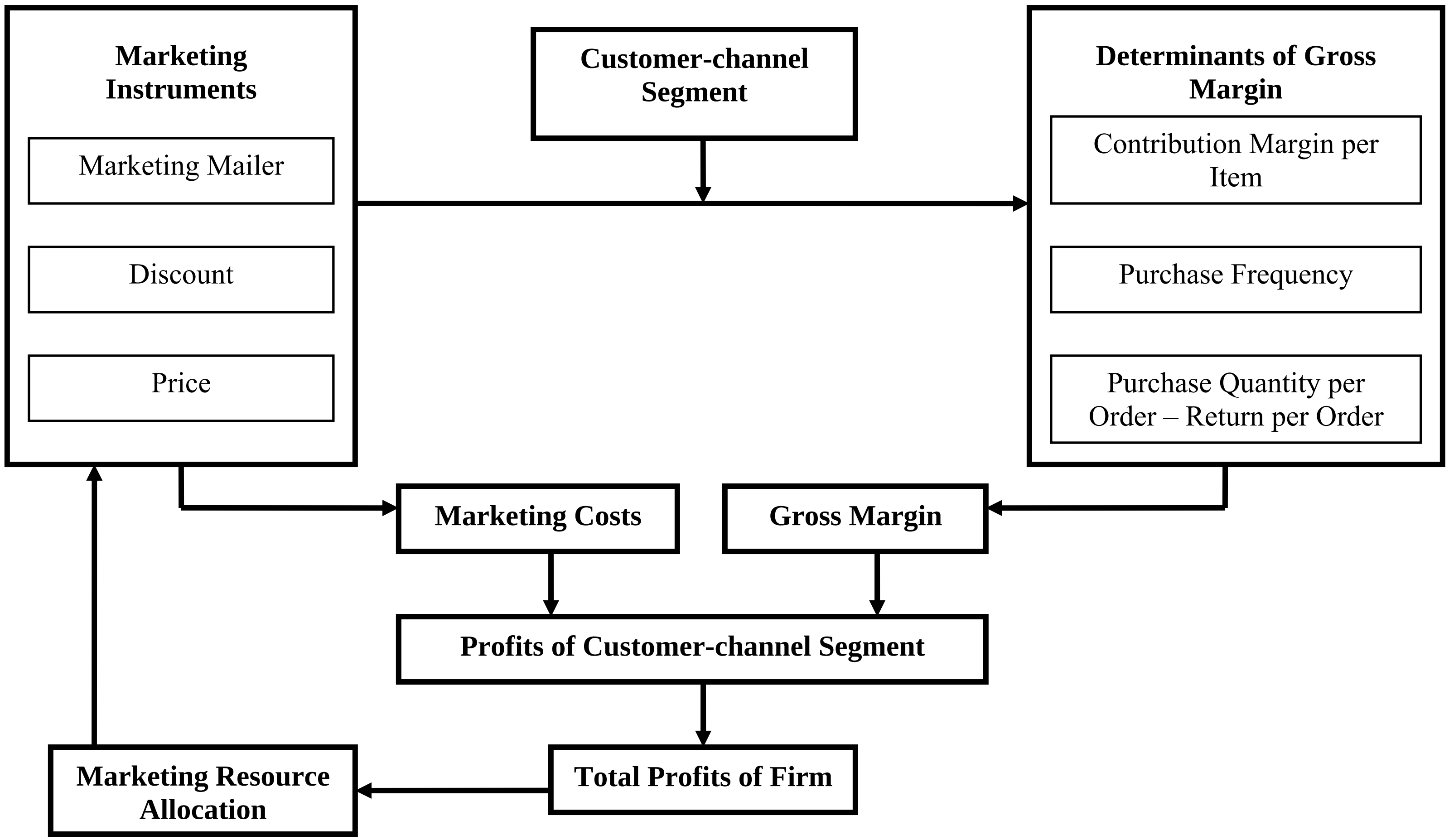


Figure 2
Profit Function

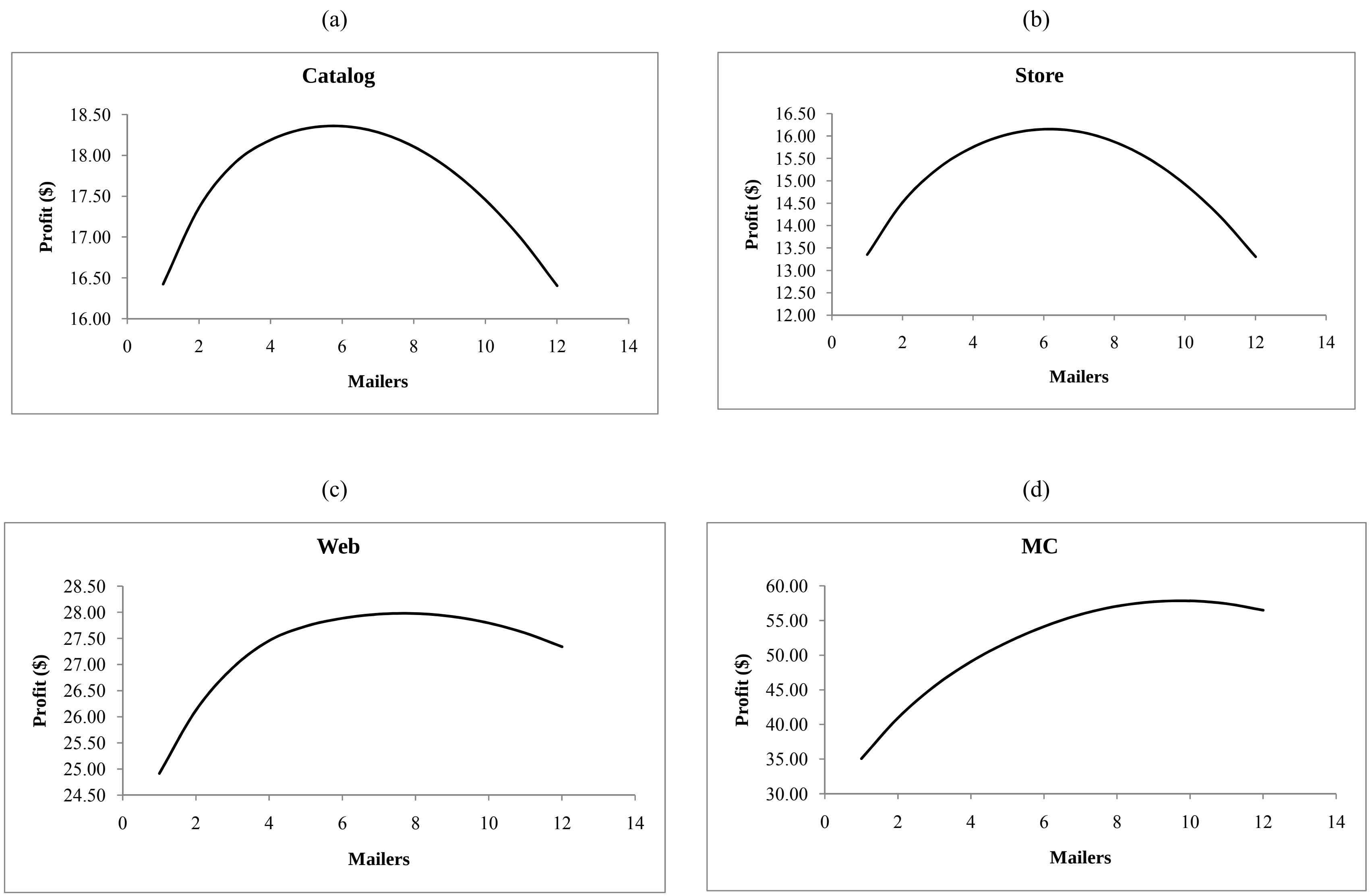
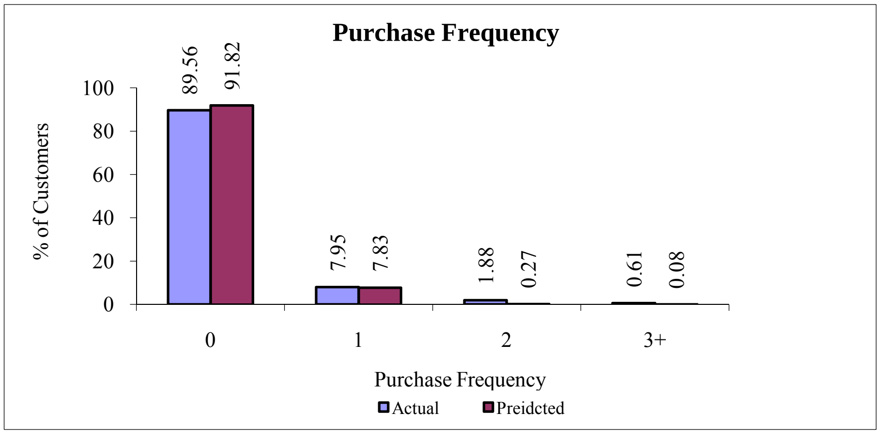
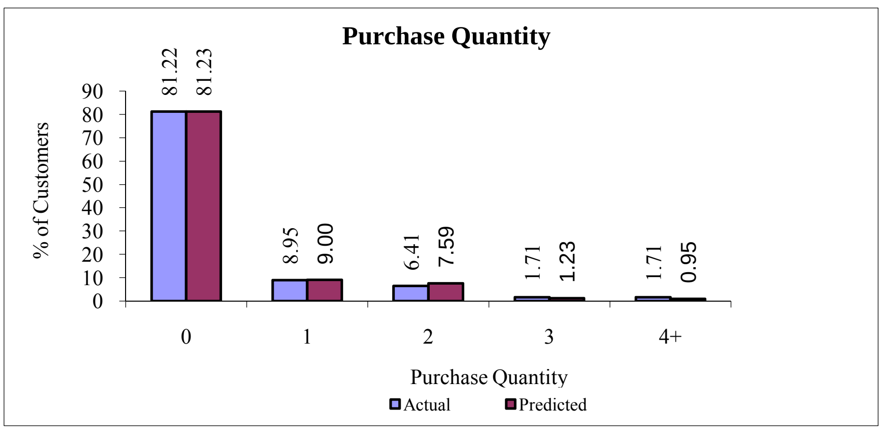


Figure 3
Predictive Validity for Catalog Only Customer-channel Segment

(a)



(b)



(c)

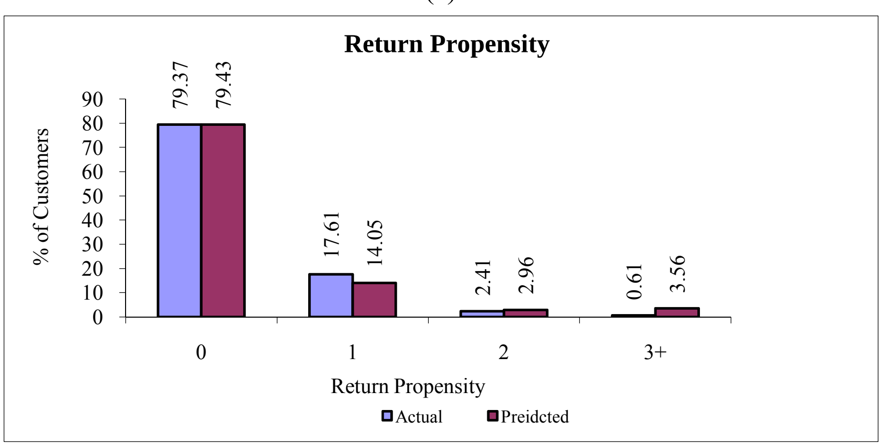
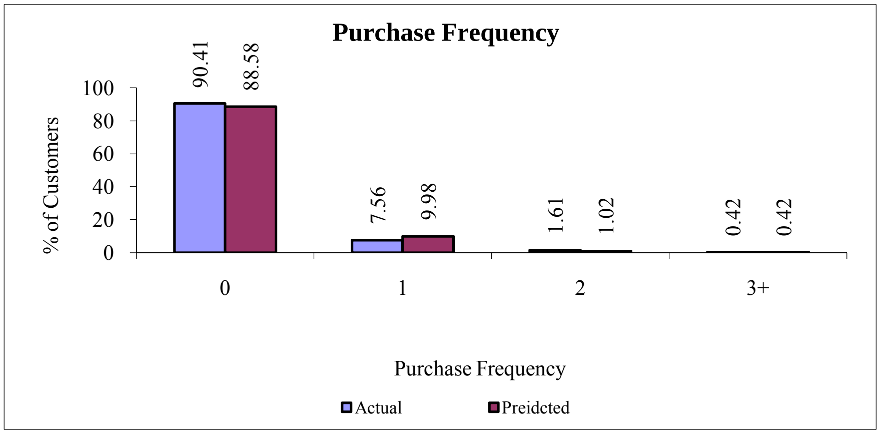
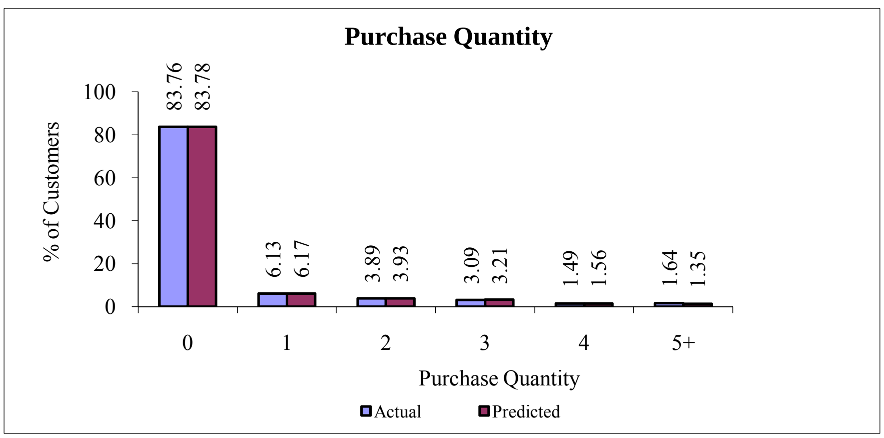


Figure 4
Predictive Validity for Store Only Customer-channel Segment

(a)



(b)



(c)

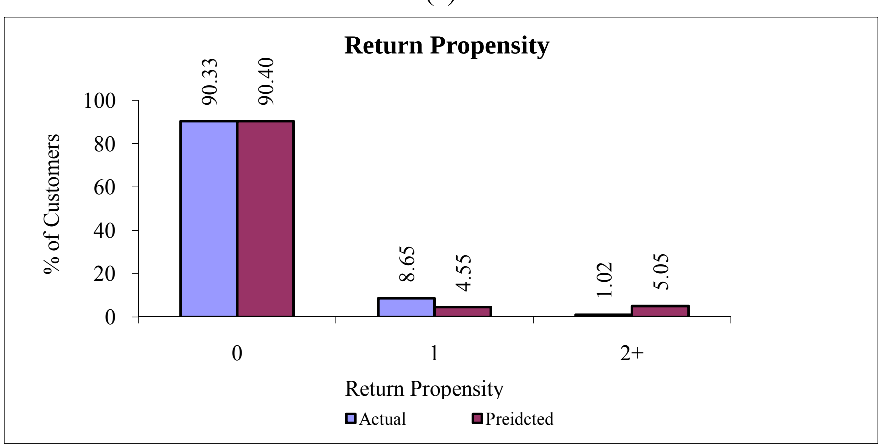
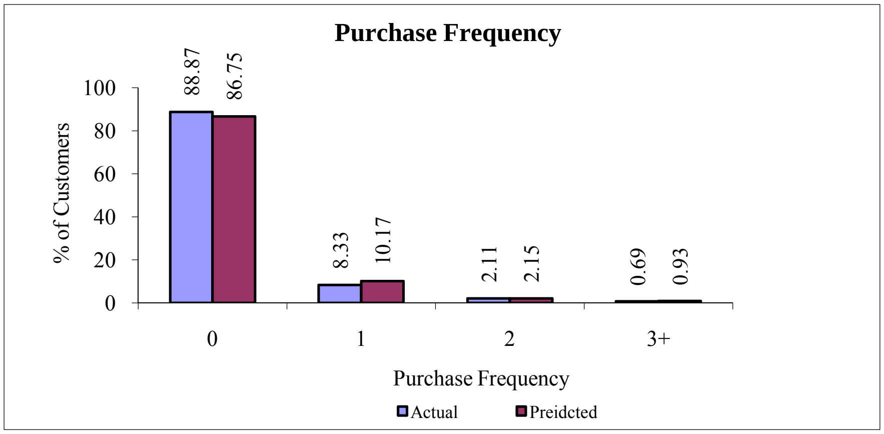
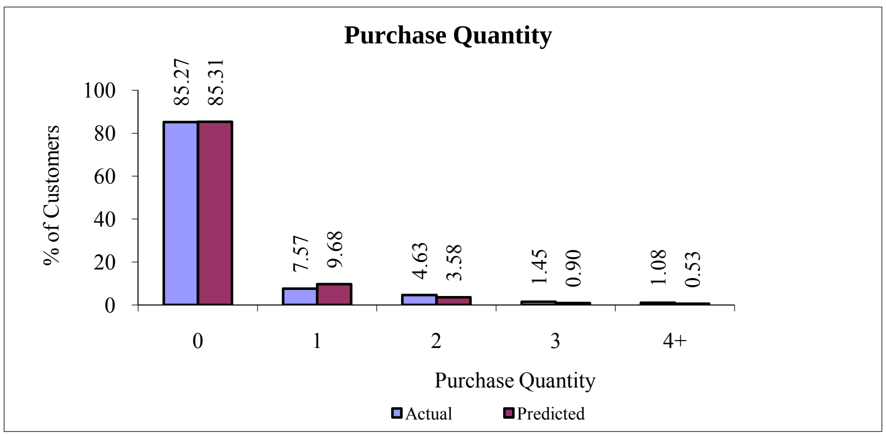


Figure 5
Predictive Validity for Web Only Customer-channel Segment

(a)



(b)



(c)

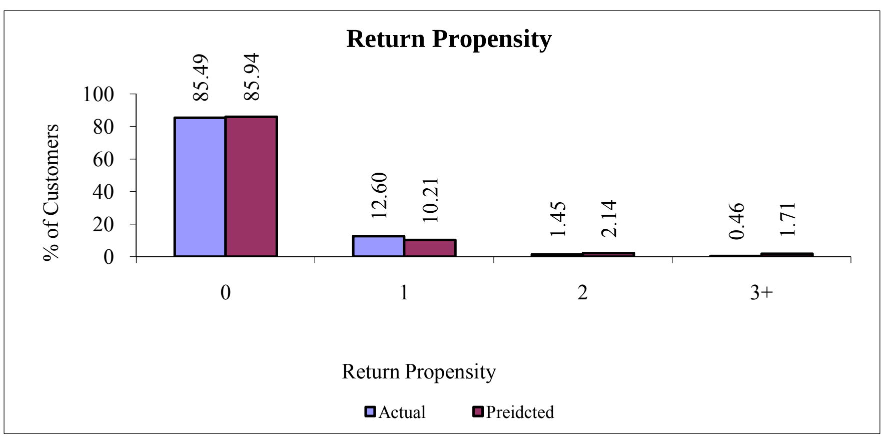
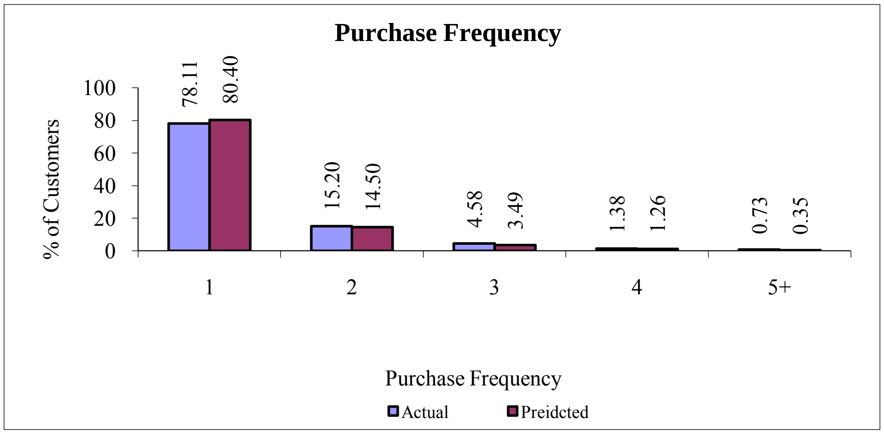
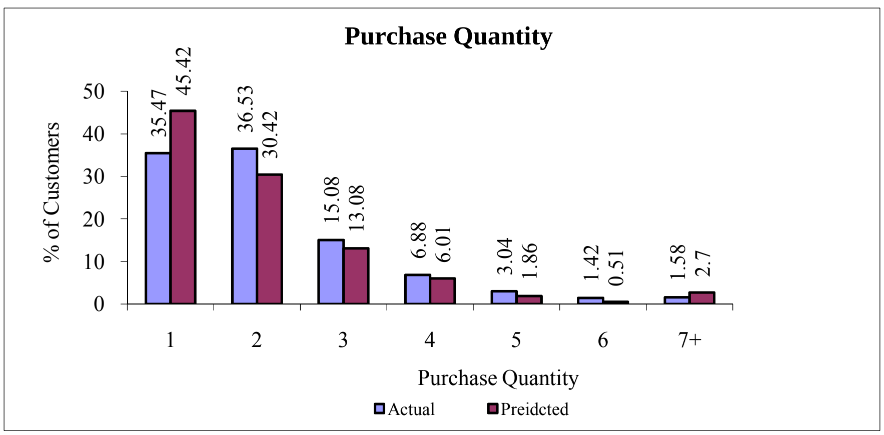


Figure 6
Predictive Validity for Multichannel Customer-channel Segment

(a)



(b)



(c)

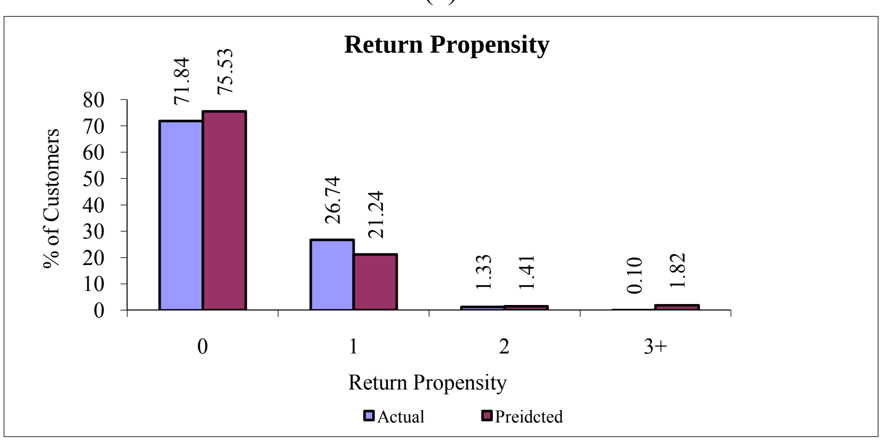
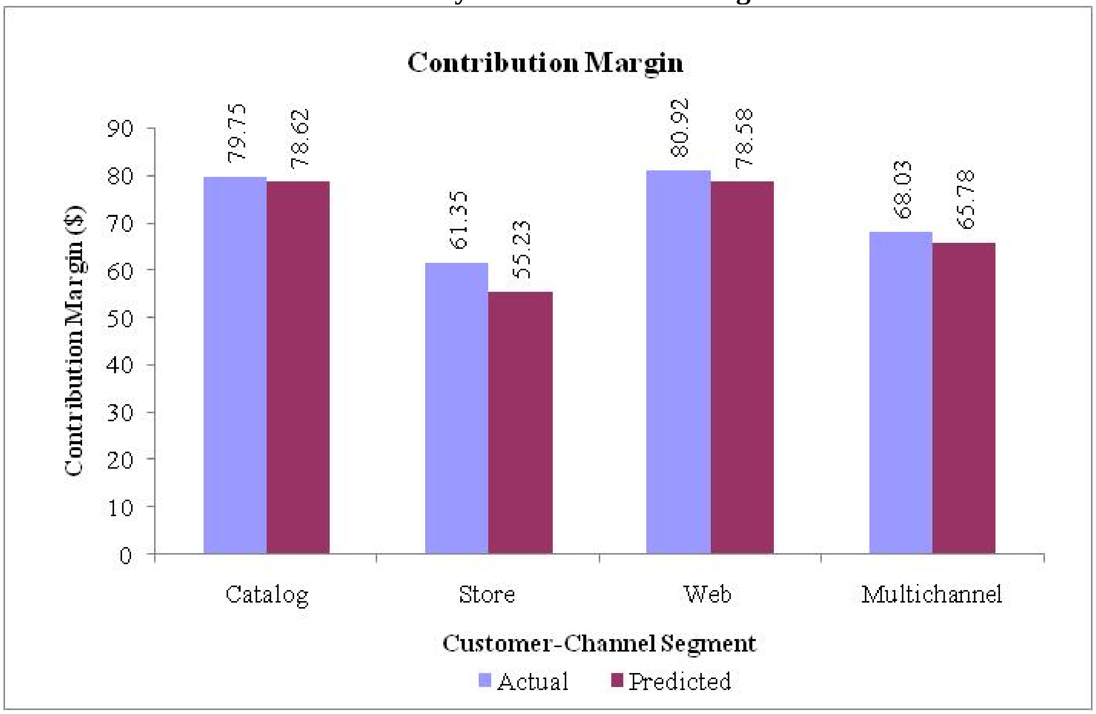


Figure 7
Predictive Validity of Contribution Margin Model



Appendix A

Resource Allocation Model

The profits derived from the customer-channel segments are the sum of the gross margins contributed by each customer belonging to the customer-channel segment k less the sum of the costs of marketing to all the customers in that customer-channel segment.

$$\Pi_k = \sum_{i=1}^{n_k} [TM_{ik}(m_{ik}) - c_m m_{ik}] \tag{A.1}$$

where TM_{ik} is the total gross margin contributed by customer i of customer-channel segment k and it is a function of m_{ik}, c_m is the unit cost of a marketing mailer which does not vary across customers, and n_k is size of customer-channel segment k.

The total margin earned from customer i of customer-channel segment k is given by:

$$TM_{ik}(m_{ik}) = NIB_{ik}(m_{ik}) \times \overline{CM}_{ik}(m_{ik}) \tag{A.2}$$

where NIB_{ik} is the net number of items bought by customer i belonging to customer-channel segment k and $\overline{CM}_{ik}$ is the average gross contribution margin per item of customer i belonging to customer-channel segment k over the j items bought by that customer $\overline{CM}_{ik}$ is given by:

$$\overline{CM}_{ik}(m_{ik}) = \frac{1}{J} \sum_{j=1}^{J} CM_{ijk}(m_{ik}) \tag{A.3}$$

CM_{ijk} is the gross contribution margin of customer i belonging to customer-channel segment k for the j^{th} item bought and is given by:

$$CM_{ijk} = P_j - C_j - D_{ijk} \tag{A.4}$$

where P_j is the price of item j, C_j is the cost of item j, and D_{ijk} is the discount offered to customer i belonging to customer-channel segment k for item j. Note that while price and cost vary across items, the promotional discount can vary across both items and customers. In other words, some customers may buy an item during a promotion period, while others may buy during a non-promotion period, leading to different contribution margins for the same item across different customers.

The net items bought by customer i belonging to customer-channel segment k is given by:

$$NIB_{ik}(m_{ik}) = [IPO_{ik}(m_{ik}) - IRPO_{ik}] \times NO_{ik}(m_{ik}) \tag{A.5}$$

where IPO_{ik} is the number of items bought per order by customer i belonging to customer-channel segment k, $IRPO_{ik}$ is the number of items returned per order by customer i belonging to customer-channel segment k, and NO_{ik} is the number of orders by customer i belonging to customer-channel segment k. The firm's profits from customer-channel segment k of size n_k customers is given by:

$$\Pi_k = \sum_{i=1}^{n_k} [(IPO_{ik}(m_{ik}) - IRPO_{ik}) \times NO_{ik}(m_{ik}) \times \overline{CM}_{ijk}(m_{ik}) - c_m m_{ik}] \qquad \text{(A.6)}$$

Using this disaggregated approach, we decompose the total margin derived from a customer i into four different sub models: purchase frequency (NO_{ik}), purchase quantity per order (IPO_{ik}), product returns per order ($IRPO_{ik}$), and gross contribution margin ($\overline{CM}_{ijk}$). This approach enables us to understand the effect of marketing efforts on order size and on up-selling.

Appendix B

Purchase Frequency Model

The underlying assumptions of the BG/NBD model are[8]:

a) A customer can become inactive immediately after her purchase. This customer drop out is distributed across transactions according to a geometric distribution,
b) the dropout rate across customers is beta distributed with shape parameters ‘a’ and ‘b’,
c) while active, a customer’s transaction rate follows Poisson distribution, and
d) the transaction rate across customers is gamma distributed with shape parameter ‘r’ and scale parameter α .

The model requires three pieces of information from each individual customer: the purchase frequency or the number of repeat transactions (x), the time since the first purchase *(T)*, and the time of last purchase (t_x). While the customer dropout process is inherently a stochastic process, the transaction rate of customers who are active can be modeled for the effect of covariates. The distribution function for purchase frequency is given by:

$$g(x, t_x, T \mid r, \alpha, a, b) = \frac{\Gamma(b+x)\Gamma(a+b)}{\Gamma(a+b+x)\Gamma(b)} \frac{\Gamma(r+x)\alpha^r}{\Gamma(r)\Gamma(\alpha+T)^{r+x}} + \delta_{x>0} \frac{\Gamma(a)\Gamma(b)\Gamma(a+1)\Gamma(b+x-1)}{\Gamma(a+b)\Gamma(a+b+x)} \frac{\Gamma(r+x)\alpha^r}{\Gamma(r)\Gamma(\alpha+t_x)^{r+x}} \ , \qquad \text{(B.1)}$$

where, Γ is the gamma function. We introduce covariates to the model as a function of the shape parameter of gamma heterogeneity in following form.

$$\alpha = \frac{\exp(\beta z^{'})}{r} \ , \qquad \text{(B.2)}$$

where $z^{'}$ is the vector of covariates and β is a vector of response parameters.

The modified log likelihood function for customer-channel segment k is given by:

$$\mathrm{LL_k(r,a,b,\beta \mid X = x_{ik}, t_{xik}, T_{ik}) =}$$

[8] We test these assumptions in our empirical analysis.

$$\sum_{i=1}^{n_k} \mathrm{Ln}\left[\frac{\left[\dfrac{\Gamma(r+x_{ik})\left(\dfrac{\exp(\beta z_{ik}^{'})}{r}\right)^{r}}{\Gamma(r)}\right]}{\left[\dfrac{\Gamma(a+b)\Gamma(b+x_{ik})}{\Gamma(b)\Gamma(a+b+x_{ik})}\right]}\left[\left(\frac{1}{\dfrac{\exp(\beta z_{ik}^{'})}{r}+T_{ik}}\right)^{r+x_{ik}} + \delta_{x>0}\left(\frac{a}{b+x_{ik}-1}\right)\left(\frac{1}{\dfrac{\exp(\beta z_{ik}^{'})}{r}+t_{x\,ik}}\right)^{r+x_{ik}}\right]\right] \tag{B.3}$$

where x_{ik}, t_{xik}, and T_{ik} are the purchase frequency, time of last purchase, and time since first purchase of customer i of customer-channel segment k. The shape parameters of the joint distribution are ‘r’, ‘a’, and ‘b’.

The conditional expectation of purchase frequency of a customer in a non-overlapping prediction time window ‘t’, given the shape parameters, the response parameters, and the observed purchase behavior of the same customer in the estimation window of length ‘T’ is given by: (see Fader, Hardie, and Lee 2005a for details)

$$E(Y(t)\,|\,r,a,b,\beta,X=x,t_x,T)=$$

$$\frac{\dfrac{a+b+x-1}{a-1}\left[1-\left(\dfrac{\dfrac{\exp(\beta z^{''})}{r}+T}{\dfrac{\exp(\beta z^{''})}{r}+T+t}\right)^{r+x} {}_2F_1\left(r+x,b+x;a+b+x-1;\dfrac{t}{\dfrac{\exp(\beta z^{''})}{r}+T+t}\right)\right]}{1+\delta_{x>0}\dfrac{a}{b+x-1}\left(\dfrac{\dfrac{\exp(\beta z^{''})}{r}+T}{\dfrac{\exp(\beta z^{''})}{r}+t_x}\right)^{r+x}} \tag{B.4}$$

where ‘Y(t)’ is the expected number of transactions for a given customer in a future time window ‘t’, given her purchase behavior in a non-overlapping previous time window ‘T’ and $z^{''}$ are value of covariates for the given individual in the prediction window. ${}_2F_1$ is a Gauss Hypergeometric function (sum of a hypergeometric series) with two parameters of Type 1 and one parameter of Type 2. The expected number of transactions given in Equation (B.4) incorporates the probability of a customer remaining active for the time duration ‘t,’ multiplied by the expected transaction rate and the time duration of the window.

The ${}_2F_1$ is Gauss Hypergeometric function is given in Equation B.5 (see Fader, Hardie, and Lee 2005c for details). Note that one hypergeometric function is evaluated for each customer in the prediction window.

$${}_2F_1(P1,P2;P3;P4)=\sum_{s=0}^{\infty} ST_s \tag{B.5}$$

where, P1 = r + x, P2 = b + x, P3 = a + b + x -1, and P4 = $\dfrac{t}{\dfrac{\exp(\beta z^{''})}{r} + T + t}$ are four parameters of the hypergeometric series as given in Equation (B.5). ST_s denotes the s^{th} term of the hypergeometric series which is given by:

$$ST_s = \frac{(P1)_s (P2)_s (P4)^S}{(P3)_s s!} \tag{B.6}$$

where $(P1)_s$ is the ascending factorial given by (P1)(P1+1)(P1+2)…..(P1+s-1). The other two terms, $(P2)_s$ and $(P3)_s$ are defined in a similar fashion.

Purchase Quantity Model

The NBD model assumes that purchase quantity per order is Poisson distributed with gamma heterogeneity across customers with shape parameter 'c' and scale parameter ϕ. The distribution function for purchase quantity is given by:

$$l(f \mid c, \phi) = \frac{\Gamma(f + \phi^{-1})}{\Gamma(f+1)\Gamma(\phi^{-1})} \left(\frac{c}{\phi^{-1} + c} \right)^f \left(\frac{\phi^{-1}}{\phi^{-1} + c} \right)^{\phi^{-1}}, \tag{B.7}$$

where f is the purchase quantity and Γ is the gamma function. We introduce the covariates in the model as a function of the scale parameter that follows a gamma distribution as given below.

$$\phi = \frac{\exp(\eta z^{'})}{c}, \tag{B.8}$$

where η is the response parameter of covariates $z^{'}$.

The likelihood function of the NBD model for customer-channel segment k is given by

$$LL_k(c, \eta \mid F = f_{ik}) = \sum_{i=1}^{n_k} Ln\left[\frac{\Gamma(f_{ik} + c)}{\Gamma(f_{ik}+1)\Gamma(c)} \frac{(\eta z^{'}_{ik})^c (c)^{f_{ik}}}{(\exp(\eta z^{'}_{ik}) + c)(c + f_{ik})} \right] \tag{B.9}$$

where f_{ik} is the purchase quantity per order of customer i from customer-channel segment k.
The purchase quantity per order of a customer is not independent of her purchase frequency. For example, consumers who order more frequently may have smaller order sizes and vice-versa. We correct for the dependence by conditioning the expected purchase quantity per order of a given customer on her purchase frequency (Morrison and Schmittlein 1988). The conditional expectation of purchase quantity per order for consumers with at least one repeat purchase is given by:

$$E(G(y) \mid Y = y > 0; c, f, \eta, x, y) = \frac{(c+f)y}{\dfrac{\exp(\eta z^{''})}{c} + x} \tag{B.10}$$

The conditional expectation of purchase quantity per order for consumers with no repeat purchase is

$$E(G(y)\mid Y=0;c,\eta,x,y)=\frac{(1-\gamma)\left(\dfrac{cy}{\dfrac{\exp(\eta z^{"})}{c}+x}\right)\left(\dfrac{\dfrac{\exp(\eta z^{"})}{r}}{\dfrac{\exp(\eta z^{"})}{c}+x}\right)^{c}}{\gamma+(1-\gamma)\left(\dfrac{\dfrac{\exp(\eta z^{"})}{c}}{\dfrac{\exp(\eta z^{"})}{c}+x}\right)^{c}} \tag{B.11}$$

where G(y) is the conditional expectation of purchase quantity per order for a customer with purchase frequency 'y' in the prediction window, 'x' is the purchase frequency of the customer in the estimation window, 'f' is the average purchase quantity per order in estimation window, $z^{"}$ are the values of covariates in the prediction window, and γ is the proportion of consumers in the dataset with no repeat purchase.

Product Return Propensity Model

We assume return per order to be Poisson distributed with gamma heterogeneity, having a shape parameter 'd' and a scale parameter μ. The distribution function for return propensity is given by:

$$s(h\mid d,\mu)=\frac{\Gamma(h+\mu^{-1})}{\Gamma(h+1)\Gamma(\mu^{-1})}\left(\frac{d}{\mu^{-1}+d}\right)^{h}\left(\frac{\mu^{-1}}{\mu^{-1}+d}\right)^{\mu^{-1}} \quad , \tag{B.12}$$

where h is the number of items returned per order by the customer and Γ is the gamma function.

The likelihood function for the NBD model for customer-channel segment k without covariates is very similar to the likelihood function in Equation (B.13).

$$LL_k(d,\mu\mid H=h_{ik})=\sum_{i=1}^{n_k}\mathrm{Ln}\left[\frac{\Gamma(h_{ik}+d)}{\Gamma\left(\dfrac{h_{ik}+1}{d}\right)}\left(\frac{\mu}{\mu+1}\right)^{d}(\mu+1)^{h_{ik}}\right] \tag{B.13}$$

where h_{ik} is the number of returns per order for a given customer i from customer-channel segment k. Return per order of a customer is related to her purchase behavior and hence we do not assume independence of these processes. Customers who order more often are more likely to return a product because of several stochastic unobserved processes as wrong items, wrong fits, defective units, and unsatisfactory product evaluation. To correct for this dependence bias, we condition the expected return per order in the prediction window on past purchase frequency and expected order frequency. The conditional expectation of the number of returns per order for consumers with at least one repeat purchase is given by:

$$E(I(y)\mid Y=y>0;d,h,\mu,x,y)=\frac{(d+h)y}{\mu+x} \tag{B.14}$$

The conditional expectation of the number of returns per order for consumers with no repeat purchase is given by:

$$E(I(y) \mid Y=0; d, \mu, x, y) = \frac{(1-\gamma)\left(\frac{dy}{\mu+x}\right)\left(\frac{\mu}{\mu+x}\right)^{d}}{\gamma + (1-\gamma)\left(\frac{\mu}{\mu+x}\right)^{d}} \tag{B.15}$$

where 'I(y)' is the conditional expectation of return per order for a customer with purchase frequency 'y' in the prediction window, 'x' is the purchase frequency of the customer in the estimation window, 'h' is the average return per order in estimation window, and γ is the proportion of consumers in the dataset with no repeat purchase.

Contribution Margin Model

If a customer orders 'x' number of times in a given window, where the size of each order is 'f', the total number of items ordered by a customer are x*f. If w_{11}, w_{12}, w_{21}......., w_{lj} are the contribution margins of the f^{th} item bought by a customer on the x^{th} purchase occasion, the average contribution margin per item for that customer is given by:

$$w_{xf} = \frac{\sum_{l=1, j=1}^{x, f} w_{lj}}{x \times f} \tag{B.16}$$

Our goal is to first test the effect of marketing efforts on w_{xf} and then to predict the expected value of the contribution margin per item for a given customer in the holdout window. If we know exactly which item a customer is going to buy in a prediction window, the prediction of average contribution margin per item for that customer is not required. However, for firms marketing multiple product categories with hundreds of SKUs in each category, it is almost impossible to predict customer's exact item choice in a prediction window. Alternatively, we predict the average contribution margin per item for a given customer with the assumption that as x*f tends to infinity, w_{xf} tends to the expected value of the contribution margin per item for a given customer. Because this approach would require observation of customer behavior for an extremely long time window, we assume that there is heterogeneity in contribution margin across the items ordered by a customer and also across customers. The following are the assumptions of the Gamma-Gamma model.

a) If w_{11}, w_{12}, w_{21}......., w_{lj} are gamma distributed with shape parameter 'p' and scale parameter ν, the average value of the contribution margin per item for a given customer (w_{xf}) will be gamma distributed with shape parameter p*x*f and scale parameter ν*x*f,
b) To account for heterogeneity in the value of w_{xf} across customers, the model assumes that ν is distributed across customers according to a gamma distribution with shape parameter 'q' and scale parameter θ, and
c) If 'p' is assumed to be constant across customers, the joint marginal distribution of w_{xf} will be distributed with shape parameters 'p' and 'q', and scale parameter θ.
   The distribution function for contribution margin is given by:

$$u(x, f, w_{xf} \mid p, q, \theta) = \left(\frac{\Gamma(pxf + q)}{\Gamma(pxf)\Gamma(q)} \right) \left(\frac{\theta^q w_{xf}^{pxf-1} xf^{\ pxf}}{(\theta + w_{xf}\, xf)^{pxf+q}} \right), \tag{B.17}$$

where Γ is gamma function. We introduce covariates in this "regression to the mean" approach model as a linear form of the scale parameter θ of the joint marginal distribution in the following form.

$$\theta = \omega z', \tag{B.18}$$

where ω are the response parameters of the marketing efforts z'.

The likelihood function of the marginal distribution with the effect of covariates for customer-channel segment k is given by:

$$\mathrm{LL_k(p,q,\omega \mid w_{ikxf}, x_{ik}, f_{ik})} = \sum_{i=1}^{n_k} \mathrm{Ln}\left[\frac{\Gamma(px_{ik}f_{ik} + q)}{\Gamma(px_{ik}f_{ik}) - \Gamma(q)} \frac{(\omega z'_{ik})^q (w_{ikxf})^{(px_{ik}f_{ik}-1)} (x \times f_{ik})^{(px_{ik}f_{ik})}}{(\omega z'_{ik} + w_{ikxf} x_{ik} f_{ik})^{(px_{ik}f_{ik}+q)}} \right] \tag{B.19}$$

where 'w_{ikxf}' is the average contribution margin per item for customer i belonging to customer-channel segment k with 'x_{ik}' orders and 'f_{ik}' purchase quantity per order. To optimize marketing efforts in a future time period 't', we need to predict the conditional expected contribution margin per item of a randomly chosen individual. The conditional expectation of contribution margin per item for a customer given the model parameters, response parameters, observed contribution margin per item, purchase frequency and purchase quantity per order for the customer in estimation window is given by (see Fader, Hardie, and Lee 2005b for details):

$$\mathrm{E(U \mid p, q, \omega, x, f, w_{xf})} = \left(\frac{q-1}{Pxf + q - 1} \right) \left(\frac{\omega z'' p}{q-1} \right) + \left(\frac{pxf}{pxf + q - 1} \right) w_{xf} \tag{B.20}$$

where 'U' is the expected value of contribution margin per item for a given customer. Note that the covariates only influence the population mean (variation across customers) and not the mean value of margin per item for a given customer (variation across items for a given customer). The value of 'U' in the above equation is the weighted average of population mean and the mean value of contribution margin per item for a given customer (w_{xf}). The terms, $q-1$ and $p*x*f$ divided by their sums are the weights placed on the population mean and the mean of value of margin per item for a given customer, respectively.

Appendix C

Robustness Checks

We performed several robustness checks. First, we tested for the assumptions of the independence of each of the purchase frequency and purchase quantity models from the contribution margin model. In our data, the correlation between frequency and contribution margin (-0.01). and between quantity and contribution margin (-0.08) are low, suggesting that the assumption is reasonable.

Second, we tested for the potential endogeneity of channel choice by testing whether marketing efforts lead to a customer switching from one channel segment to another. Little over 1% of customers transitioned from one segment to another between the estimation and prediction windows. We estimated each of our models by including and excluding these transitioning customers in the data. The results and substantive findings remain unchanged.

Third, we performed the analyses using an alternative operationalization of price discount—an absolute measure of price discount. The results were substantively similar.

Fourth, to examine if the responsiveness to marketing mailers differed by the type of marketing mailer, we estimated a set of models in which, we allowed for two types of marketing mailers: regular marketing mailer and promotional mailer. We did not find a substantial difference in customers' responsiveness to the two types of mailers.

Fifth, we ran our optimization model under the scenario of fixed budget and compared the allocation results to those when the budget is allowed to vary. Under a fixed budget scenario, by reallocating marketing resources between segments, the profits can be improved by only 7.35% compared to a potential increase of 33.14% when the marketing budget is allowed to vary. In other words, 26% of the increase in profit in the unconstrained budget scenario can be attributed to budget increase while the remaining 7% can be assigned to efficient reallocation.

Sixth, to examine whether the customer's geographical proximity to the nearest physical store led to different purchase patterns across the channel segments, we estimated a model in which we included the distance of the customer's home zip code from the nearest store's zip code as an additional covariate in the models. We do not find distance to closest store to be statistically significant to customers' purchase frequency, purchase quantity, and contribution margin for any of the four segments. Looking at summary statistics, we find that the average 'catalog only', 'store only', 'Web only', and 'multichannel' customer lives about 56, 25, 53, and 41 miles, respectively away from the nearest store. A cursory look at these average distances to the closest store for each of these segments suggests that distance may play a role in channel preference of customers, but as suggested by the above results, it does not play role in determining customers' purchase behavior.

Finally, to check if the multichannel customer's average purchase quantity and monetary value are high because of purchasing from more than one channel or because they are inherently high, we examined the average quantity and monetary value of all the customers during their *first* purchase. The average monetary value of the multichannel segment is the highest ($212) and its average order size (2.30) is higher than that of the single channel segments except that of the 'store only' segment (2.41). These results are consistent with those across all purchases. Therefore, it is unlikely that the multichannel segment is more valuable merely because its customers buy from more than one channel.