Marketing Science Institute Working Paper Series 2012
Report No. 12-107

The Role of Attitudinal Information in CLV-based Customer Management

Rajkumar Venkatesan, Werner Reinartz, and Nalini Ravishanker

Report Summary

The ability to build and maintain extensive behavioral databases about customers has prompted a number of firms to de-emphasize "soft" attitudinal customer information. Proponents of behavioral data argue that customer purchase behavior encapsulates underlying attitudes, and because decision makers are mainly concerned with customer behavior, attitudes do not deserve much attention.

This study questions this conventional wisdom. Rajkumar Venkatesan, Werner Reinartz, and Nalini Ravishanker determine how the inclusion of customers' attitudes affects predictions of customer lifetime value and thus a firm's customer management strategy. They evaluate which aspects of customer behavior – retention or sales – depend more on customer attitudes, and whether knowledge of customer attitudes is more important for managing certain customers.

On the basis of monthly sales calls, sales, and survey-based attitude information collected from customers of a multinational pharmaceutical firm over three years, the authors develop a zero-inflated Poisson model framework to model both retention and sales simultaneously. The hierarchical structure of the model allows for the influence of attitudes on retention and sales.

The results show that information on customer attitudes can help companies both explain and predict customer responses better (sales and retention). The positive effect of better attitudes on lifetime value works predominantly through the retention aspect. This implies that firms should consider customer retention rates rather than sales for any campaign focused on improving the customer's emotional attachment.

Furthermore, the study suggests that the incremental profits from including customer attitudes in lifetime value and resource allocation models are greatest for mid-tier customers. These customers' attitudes provide a forward-looking measure that can effectively discriminate mid-tier customers with the potential to grow from those whose profitability is likely to decrease. The study therefore suggests that firms may be actually overspending on top-tier customers. When customer attitude information is available, firms can improve the ROI of their CRM campaigns by balancing resources between top- and mid-tier customers.

Overall, in this study's context, the projected incremental profits from a customer management strategy informed by customer attitudes exceed the investment required to collect customer attitudes. The authors therefore encourage firms to explore avenues for measuring customer attitudes and including this information in their customer management strategy.

Rajkumar Venkatesan is Bank of America Research Associate Professor of Business Administration, Darden Graduate School of Business, University of Virginia. Werner Reinartz is Professor of Marketing, University of Cologne. Nalini Ravishanker is a Professor and the Undergraduate Program Director in Statistics, University of Connecticut.

Acknowledgments
The authors thank a multinational pharmaceutical firm for sharing the customer data used in this study. They also thank participants at the 2009 Joint Statistical Meeting Conference and the 2010 Marketing Science Conference; seminar participants at the University of Groningen and

University of Maryland; and Sunil Gupta, Peter Verhoef, Pete Fader, Harald van Heerde, Casey Lichtendahl, V. Kumar, Giandomenico Sarolli, Partha Krishnamurthy, Ian Skurnik, Paul Farris, Ron Wilcox, and Prashant Malaviya for their comments on previous versions of this manuscript.

THE ROLE OF ATTITUDINAL INFORMATION IN CLV-BASED CUSTOMER MANAGEMENT

Introduction

The ability to build and maintain extensive behavioral databases about customers has prompted a de facto de-emphasis of "soft" attitudinal information by many organizations. Proponents of behavioral data approaches argue that customer purchase behavior encapsulates underlying attitudes, and because decision makers are mainly concerned with customer behavior, there really is no need to worry about underlying attitudes. Reflective of this belief, we note a massive shift among marketing strategy researchers, toward analyses and models of customer behavior as a core dependent variable and away from the attitudinal and intention metrics that were more prevalent before the availability of customer databases (Gupta and Zeithaml 2006). Even the debate about marketing's return on investment (ROI) (Srinivasan and Hanssens 2009) reflects the underlying sentiment that attitudinal insights are insufficient at senior decision-making levels, and behavioral insights represent today's benchmarks.

Information about customer attitudes plays a role in an array of marketing contexts, spanning the domains of customer insight, advertising, and consumer decision making. With this study, which we position within the customer relationship management (CRM) domain, we question the conventional wisdom that suggests ignoring information about customer attitudes when information on individual customer behavior is available. Recent developments among practitioners indicate a renewed interest in the explicit incorporation of customer attitudes into CRM decisions. Developing marketing mix and resource allocation strategies that respond to shifts in underlying customer attitudes—*in addition to their behavioral responses*—has reemerged as a priority (Kerin and Regan 2008). For example, Siebel, dunnhumby USA, Yankelovich, ZS Associates, and IBM all have made large-scale investments in systems to track various customer attitudes,[1] information about customer–firm interactions, and customer responses. The relevance of this issue has been reinforced by the Marketing Science Institute's recognition of "new approaches to generating customer insights" and "accountability and ROI of marketing expenditures" as top research priorities.

[1] Existing CRM literature uses the umbrella term "customer attitudes" to refer to various evaluative judgments and related beliefs, norms, and perceptions, such as satisfaction (or its components), commitment, perceived fairness, or relationship quality. For our study, we use "customer attitudes" to denote customers' cognitive appraisals of product performance and salesperson performance, credibility, and knowledge.

Predicting customer lifetime value (CLV) and CLV-based customer management are the two central issues in CRM (Reinartz and Venkatesan 2008). The role of attitudinal information in those activities is yet unknown but according to the resurfacing interest very worthwhile to explore. Three topics in this context motivate our study.

Predicting CLV

Customer lifetime value (CLV) prediction is critical for (re)directing customer-level resources. Prior literature has featured an array of increasingly sophisticated CLV models with reasonable precision (Bijmolt et al. 2010). But none of them include customer attitudes—which seems surprising, considering the managerial interest in the potential predictive capability of attitudinal information and the widespread availability of such data. The incremental contribution of attitudes in predicting CLV also is a nontrivial issue, in that including more variables does not necessarily improve the predictive power of models, and attitudes normally have been measured with error in primary research, which has reduced their ability to improve CLV predictions.

CLV based customer management

Besides better CLV predictions, we also want to understand which aspects of customer behavior within the CLV formulation, i.e., the cash flow (or sales) and retention components, are most affected by customer attitudes. In addition to extending theory, such knowledge about which aspect of customer behavior is affected more by customer attitudes has important practical implications. For example, firms normally run different campaigns to prevent churn (retain customers) and to improve customer spending. If, for example, attitudes affect customer retention more than spending, customers' emotions and the nontangible benefits provided by a firm are more critical for retention than for sales. Moreover, the inclusion of information about customer attitudes should improve resource allocation processes across customers. Moving beyond a technically better statistical model, a new approach could highlight where and why attitudinal information might open opportunities for better resource allocation decisions. In particular, we are interested in effects across customers in different value tiers, an issue that has attracted keen research interest (Drèze and Nunes 2009; Rust et al. 2004).

Returns from including customer attitudes

In terms of justifying the cost and effort of including attitudinal information in CLV models, it is important to acknowledge that attitudinal information is not observable and can be generated only through conversations with customers. Collecting customer attitudes is a resource-intensive process, requiring substantial financial, information technology, and human investments. Firms are interested in the expected returns from including customer attitudes in CLV models, so that they can justify the investments required to collect that information. Furthermore, in practice, attitudinal information is available only for a limited set of customers. If CLV models were applied across the customer base, attitudinal information would need to be imputed for customers about whom information is unavailable. Associated CLV models should account for this required imputation. Nor is it clear if customer targeting decisions based on CLV predictions that use imputed or forecasted customer attitudes provide higher ROI than targeting decisions that ignore customer attitudes.

In line with these three considerations, we propose a framework to estimate the returns of including customer attitudes in CLV models, which also serves as a case example of the potential returns on including customer attitudes in CLV models. Thus, we pursue four objectives:

- For predicting customer-level CLV, does it make a difference if we include customer attitudinal information?
- If attitudinal information matters in a CLV model, for which component of CLV does that information matter most—retention or spending?
- In terms of resource allocation, does including attitudes redirect the focus to different types of customers?
- How can firms measure the incremental returns on including customer attitudes in CLV models, and do the potential benefits outweigh the costs of collecting that information?

The availability of longitudinal data about customer's attitudes and behavior and firm marketing actions enables us to assess those effects empirically. We test our conceptual framework with data from a multinational pharmaceutical company—physician prescriptions, sales calls directed toward the physicians, and survey data about the physician attitudes—over a three-year period. Similar to most other research studies in this context, we treat physicians as customers of the pharmaceutical firm, so sales calls directed toward them constitute CRM actions.

In this study, we make three substantive contributions. First, we document that including information about customer attitudes offers significant incremental predictive power in CLV models. This finding holds for both retention and sales components, though attitudes affect CLV mainly through their effect on retention, which might have important implications for the metrics and key performance indicators that firms use to diagnose their customer base. Second, a key result from the selection and resource allocation exercises are that the incremental profits from including customer attitudes are highest for mid-tier customers. Including attitude information in CLV models is most diagnostic for this group and yields the highest benefits, likely because these customer attitudes provide a forward-looking measure that can effectively discriminate mid-tier customers that have the potential to grow, as well as those that may further decrease in profitability or terminate the relationship. Whereas prior research focused largely on the value of top-tier customers, we show for the first time how firms can target mid-tier customers for profitable future growth opportunities. Third, we provide initial evidence that the returns from including customer attitudes in CLV models justify the data collection and integration efforts, at least in our empirical context. Finally, in terms of methodology, we develop a zero-inflated Poisson (ZIP) model to predict retention and sales simultaneously, using sales calls and customer attitudes. We propose an imputation framework to infer attitudes for customers whose attitude information is missing and thereby develop a CLV model that includes attitudes across the customer base.

In the next section, we review prior literature that links customer attitudes and behavior and provide our conceptual framework. We then explain the available data and the model structure we used to assess our conceptual framework. Subsequently, we discuss the results from the model estimation and derive some managerial implications of our findings. Finally, we identify the limitations of this study and avenues for further research.

Conceptual Background

Consumer behavior literature has firmly established that customer attitudes are integral to customer decision making (Alba et al. 1991). Attitudes represent global, relatively enduring evaluations of objects, issues, or persons, which may be based on behavioral, cognitive, or affective information and experiences, and can guide behavioral responses (Petty et al. 1991). Prior research shows that the predictive capability of choice models improves with attitude information, even if it is available only at a single point in time (Horsky et al. 2006) and attitudes toward both the firm and the competition should be included in sales response models (Srinivasan et al. 2010).

Even with these recommendations, important work remains to be done. Srinivasan et al. (2010) evaluate market-level changes in attitudes toward the brands in a category, using sales of corresponding brands in that category. Rust et al. (2004) use customer surveys to link a firm's marketing inputs, sales, acquisition and retention rates, and customer lifetime value, which means that they mainly compare the importance of different marketing investments for various firms within an industry.

In our study, we instead link typical behavioral data available in CRM systems, such as sales and sales calls, to customer attitudes obtained from surveys. In so doing, we build on previous efforts and extend them to a CRM context in which marketing actions (i.e., sales calls) vary substantially across customers. We propose and empirically evaluate a chain of effects model (Figure 1) of the mechanism by which customer attitudes and CRM activities interact to influence sales and CLV at the individual customer level. (Figures and tables follow References.)

In our model, customer retention and sales depend on sales calls and the customer's past experience with the product. Customer attitudes should directly affect retention and sales and also moderate the relationship among sales calls, past experience, and customer behavior. We do not simultaneously model the determinants of customers' attitudes (e.g., sales calls, past sales) in our framework though, because we are interested in the cross-sectional differences in attitudes across customers, not the dynamics of changes in attitudes on customer behavior.[2]

Data

We empirically assess the proposed conceptual framework using the CRM database of a multinational pharmaceutical firm. We consider a newer drug launched by the pharmaceutical firm in a large therapeutic drug category (one of the ten largest in the United States), which we refer to as the "own drug." The therapeutic category contains multiple major drug candidates,

[2] We simultaneously modeled attitudes and used predicted instead of observed attitudes (similar to two-stage least squares) for retention and sales; the substantive results were the same. The results are available on request.

and the own drug possesses an intermediate market share. For confidentiality reasons, we cannot reveal any other information about the drug category or the pharmaceutical firm.

The database contains monthly prescription history from 6,249 physicians (customers)[3] of the own drug for 45 continuous months. The time window starts after one year after the introduction of the focal drug, which occurred in the past decade. To measure unit sales, we count the number of new prescriptions per month for a customer (physician) of the own drug. In each month, the database also provides the number of sales calls per customer (or detailing in the pharmaceutical industry) that the firm used to promote just the own drug. During these 45 months, the pharmaceutical firm also collected information about customer attitudes toward all drugs in the therapeutic category and their salespeople, with the primary goal of gathering customer feedback to inform its sales force evaluation and training. The firm did not use this customer attitude information to determine the level of sales calls for individual customers.

Each month, all customers received an e-mail, asking them to provide their responses for all the drugs in the therapeutic category. The sampling frame for the survey included the list of customers available in the firm's database, combined with the American Medical Association database, which should cover at least 95% of the physicians in the United States. The information in the survey relevant for our study included (1) customer ratings (seven-point scale) of the salesperson of each drug in terms of overall performance, credibility, knowledge of the disease, and knowledge of medications; (2) customer ratings (seven-point scale) of each drug with regard to its overall performance; (3) the customer's specialty, and (4) customer estimates of the number salesperson visits for each drug in the most recent month.

The response rate to the survey was approximately 15% of customers contacted; however, that rate increased among customers who had prescribed the own drug at least once to 35%. These statistics are similar to the response rates obtained in other studies in the pharmaceutical industry (e.g., Ahearne et al. 2007). In addition, at least one response about either the firm's own or competitors' drugs was available from 6,249 customers. More than 3,000 customers responded at least once regarding both the focal firm and its competition.

The mean level of monthly sales for own and competitive drugs and mean sales calls from the firm were not significantly different between respondents and non-respondents. Nor were there any significant differences in monthly sales, sales calls from the firm, or attitudes

[3] As we noted, we treat physicians as customers, even though they are not the end consumers of the firm's products, because customer-level CRM activities by pharmaceutical firms are directed toward physicians. We use the term customer to refer to physicians hereafter.

toward the firm's and competitors' drugs between early and late respondents. The number of survey responses for a customer had no significant correlation with attitude levels. Thus, selection bias does not appear to be an issue in our analyses; the physicians who responded to the survey were not new to the therapeutic category and did not systematically prefer the own drug.[4]

Correlations between measures of salesperson attitudes and drug attitudes ranged from .57 to .84,[5] such that customers did not discriminate much between their salesperson and drug attitudes but rather formed an overall evaluation. For each drug category, we took an average of the five attitude measures toward salespeople and drugs to measure *attitudes toward the firm*.[6]

We provide the descriptive statistics for all the variables in our model framework in Table 1. The firm obtained approximately three new prescriptions (sales) per month from each customer, and salespeople called on each customer twice a month on average. Yet we also observed large variation in the monthly level of sales per customer and the number of sales calls directed toward customers each month. On average, customers prescribed the own drug every 1.5 months. Over a three-month window, they prescribed it approximately 2.1 times, and the total number of prescriptions in this three-month window was 7.7. The p(alive) measure (Schmittlein and Morrison 1985) indicated that the empirical probability that a customer in our sample would prescribe the own drug at least once in a given month was 87%.

On a seven-point scale, customer attitudes toward the focal firm and the competition averaged 5.13 and 5.6, respectively. The within-customer difference between these attitudes was –.47, which implied that customers in our sample marginally preferred the competition. This preference also emerged from the share-of-wallet metric, which averaged 18% for the focal firm. The sample statistics reflected the population accurately; the focal firm was not the market leader in this drug category. The share-of-wallet measure can be low even when the p(alive) measure is high, because share of wallet refers to the ratio of the number of prescriptions of the own drug to the total prescriptions in the category, whereas p(alive) measures the probability that the customer prescribes at least once in a given month. Approximately 45% of the customers in our sample were specialists.

---

[4] The analysis of variance results are available on request.

[5] We measure correlations (associations) among interval scaled data, such as the attitude items, and discrete valued data, such as sales calls and sales, using Kendall's Tau, the most appropriate method. Hereafter, all mentions of "correlations" refer to Kendall's Tau.

[6] We also conducted a factor analysis of the five attitude measures, and all measures loaded on a single factor. The variance explained by the single-factor solutions for the focal firm, leaders, and challengers were .81, .84, and .82, respectively.

To measure competitor sales calls, we turned to the surveys used to collect customer attitudes. That is, the customers' recollections of the level of sales calls employed by competitors provided estimates of competitor sales calls, which matches the methods used by third-party data providers in the pharmaceutical industry, such as ImpactRx. We also evaluated the validity of these estimates by examining the correlation of customer estimates of competitor sales calls with different factors in our model. For example, customer estimates of competitor sales calls correlated negatively with sales by the focal firm ($r = -.09$), positively with the customer's attitudes toward the competition ($r = .12$), and negatively with attitudes toward the focal firm ($r = -.06$). Their estimates of the focal firm's sales calls instead showed a positive correlation ($r = .15$) with sales by the focal firm. The correlation between the level of sales calls employed by the focal firm (obtained from its sales force database) and customer estimates of its level of sales calls was greater than .50.

Model Development

We adopt an "always a share" approach to calculate CLV. That is, we assume that customers never quit a relationship and simply are dormant for extended periods if they do not buy. A customer thus is retained if he or she is active or not dormant. This assumption is suitable for the pharmaceutical industry, because a physician is unlikely to ever stop prescribing a drug completely and there is always a finite (non-zero) probability that he or she will prescribe the drug again. Therefore, the lifetime value of customer i (CLV_i) is given by:

$$CLV_i = \sum_{t=T+1}^{T+36} \frac{(1-\hat{\pi}_{it}) * \hat{y}_{it} * GM - c * \hat{D}et_{it}}{(1+d)^{t-T}}, \quad (1)$$

where

$\hat{\pi}_{it}$ = predicted probability that customer i will be dormant in time t,

$\hat{y}_{it}$ = predicted number of new prescriptions, given that customer i is retained in time t,

GM = gross margin for a single drug prescription,

c = unit cost of a sales call, and

$\hat{D}et_{it}$ = predicted number of sales calls to customer i in time t.

We use only 36 months to compute CLV, because our conversations with the focal firm revealed that it did not plan its sales force allocations for more than 36 months and considered it unlikely that market conditions would remain the same for more than three years. Furthermore, the use of discounting meant that predicted customer profits beyond three years would not substantially alter CLV calculations (Gupta et al. 2004). Total prescriptions from a customer included both new prescriptions and refills, though we used only new prescriptions to compute CLV, because sales calls had a marginal effect just on new prescriptions. Consumers or patients can obtain refills without a physician visit, and any changes to the consumers' treatment (in dosage or brand) would be recorded as a new prescription in our data.

Our model framework consisted of two parts: estimation and prediction. Our hierarchical zero-inflated Poisson (ZIP) model simultaneously includes the three major components of CLV: retention, sales, and the firm investments, i.e., sales calls. In the estimation stage, we acknowledged that observed sales calls may be endogenous to customer behavior, because firms may allocate more sales calls to customers with higher expected sales or greater responsiveness to sales calls. In the prediction part, to also acknowledge that not all customers responded to the attitude surveys, we developed an ordered Probit model to predict attitudes among customers who never responded. The estimates from the hierarchical ZIP and ordered Probit models provided the predictions of CLV.

Zero-inflated Poisson model

In each month t (t = 1 to T), we observed the level of sales (number of new prescriptions) for a customer i (i = 1 to N) and the level of sales calls (Det_{it}) directed toward that customer. We assumed that sales (y_{it}) from customer i in month t would follow a ZIP model (Lambert 1992), such that customer i in time t can belong to either of two latent (unobserved) states, dormant or inactive, B_{it}=1, versus active, B_{it}= 0. The related interpretation suggests a high probability of no new prescriptions from customers in the dormant state. When the customer is active, the number of new prescriptions can assume values of k = 0, 1, 2, …. Market forces, marketing, and other influences likely push customers from the active to the dormant state, and vice versa. As mentioned, we assume a customer never quits the relationship, such that there is always a finite probability ($1 - \pi_{it}$) that the customer will return to prescribing the firm's drugs. Under the ZIP model, the probability that sales (y_{it}) from customer i in month t equal to k is:

$$p(y_{it} = 0 \mid \lambda_{it}, \pi_{it}) = \pi_{it} + (1 - \pi_{it}) \exp(-\lambda_{it})$$

$$p(y_{it} = k \mid \lambda_{it}, \pi_{it}) = (1 - \pi_{it}) \frac{\lambda_{it}^{k} \exp(-\lambda_{it})}{k!}, \ k = 1, 2, \ldots, \quad (2a)$$

where $\lambda_{it} > 0$.

From Equation 2a, we learn that customer i is active, $\pi_{it} = 0$, when sales reach at least one new prescription in time t (i.e., $y_{it} > 0$). Customer i could belong to the active state with probability π_{it} or the dormant state with probability $1 - \pi_{it}$ in time t when we do not observe sales, or $y_{it} = 0$. We therefore include the term $(1 - \pi_{it}) \exp(-\lambda_{it})$ when modeling the probability that sales equal 0, or p(y = 0). Our formulation is a special case of the hidden Markov model structure (Netzer et al. 2008) with two states, active and dormant. Both λ_{it} and π_{it} are unknown customer-specific parameters, modeled as functions of observed covariates (Ghosh et al. 2006). For Figure 2, we rewrite Equation 2a as a mixture model of latent random variables, V_{it} and B_{it}:

$$y_{it} = V_{it}(1 - B_{it}), \tag{2b}$$

$B_{it} \sim \text{Bernoulli}(\pi_{it})$, and

$V_{it} \sim \text{Poisson}(\lambda_{it})$.

The expected number of new prescriptions from physician i in time t, which represents the Poisson mean λ_{it}, likely is a function of the number of sales calls directed toward customer i in month t (Det_{it}). We accommodate the diminishing returns of sales calls on sales (Reinartz and Venkatesan 2008) by including the log of the number of sales calls in each month in the regression function for λ_{it}.[7] Furthermore, we use the time since the last prescription ($Recency_{it}$), number of months with positive sales ($Frequency_{it}$), and cumulative level of sales (Monetary $Value_{it}$) to capture the customer's relationship history with or behavioral loyalty to the firm (Rust and Verhoef 2005). Recency, frequency, and monetary value are common metrics in CRM literature to capture past customer value. With monetary value, we also can capture the carryover effect of sales calls on future sales. We calculate these metrics for time t as moving averages during the three months prior to t, that is, t – 1, t – 2, and t – 3. The natural logarithm of the mean λ_{it} is therefore modeled as:

$$\ln(\lambda_{it}) = \beta^{\lambda}_{0i} + \beta^{\lambda}_{1i}\ln(Det_{it} + d) + \beta^{\lambda}_{2i}Recency_{it} + \beta^{\lambda}_{3i}Frequency_{it} + \beta^{\lambda}_{4i}MonetaryValue_{it} \tag{3}$$

We also add d to the level of sales calls to remove any concerns about 0 values. Specifically, d = 1 is the smallest number that will not create large outliers in the distribution of $\ln(Det_{it} + d)$. We

[7] We also considered a model with linear and squared terms and the coefficient of the quadratic term for sales calls was not significantly different from 0. Perhaps sales calls lack sufficient variation to capture the entire range of the inverted U-shaped relationship, or, such a relationship does not exist in our data.

model the probability of the Bernoulli random variable, B_i (Equation 2b), which represents the probability that a customer is inactive in time t, as:

$$\begin{aligned} \log\mathrm{it}(\pi_{it}) = {} & \beta^{\pi}_{0i} + \beta^{\pi}_{1i}\ln(\mathrm{Det}_{it} + d) + \beta^{\pi}_{2i}\mathrm{Recency}_{it} + \beta^{\pi}_{3i}\mathrm{Frequency}_{it} \\ & + \beta^{\pi}_{4i}\mathrm{MonetaryValue}_{it} + \beta^{\pi}_{5i}\,p(\mathrm{Alive})_{it-1} \end{aligned} \tag{4}$$

In addition to the variables that affect the mean of the Poisson distribution of sales (λ_{it}), we include the empirical probability that a customer was expected to be active in time t – 1, $p(\text{alive})_{it-1}$, to predict the probability that a customer is inactive in time t (π_{it}). Following Schmittlein and Morrison (1985), we obtain the probability that a customer is still alive or active with the firm at time t as

$$p(\text{Alive})_{it} = (lt/t)^{n}, \tag{5}$$

where, lt is the time of the last customer purchase from the firm before t, and n is the number of purchases by the customer from the firm from t = 1 to lt.

Because the p(alive) metric considers both interpurchase times, through (lt/t), and the frequency of customer purchases, n, it can distinguish between long periods of customer dormancy and churn. The lagged expectation of a customer's activity status, $p(\text{alive})_{it-1}$ not only provides a measure of state dependence in the customer's activity status but also enables us to differentiate a customer's activity status and sales conceptually at time t.[8]

Hierarchical model

With a hierarchical model of customer-specific coefficients, $\beta_i = (\beta^{\lambda}_i, \beta^{\pi}_i)$, we can assess the influence of customer attitudes and observed customer heterogeneity on sales and customer retention. Differences across firm and competitive metrics, rather than levels of marketing inputs and customer mindset metrics, generally are recognized as better predictors of customer behavior (Gatigon and Xuereb 1997; Mizik and Jacobson 2009; Stahl et al. 2010). Therefore, we include the difference between average customer attitudes toward the firm and its competitors, which entails relative customer attitudes toward the firm ($\overline{\mathrm{rela}}_i$), in the hierarchical model. Some customers may provide higher (or lower) ratings on all the survey questions, but the difference in their responses to different items in the survey cannot be biased by their response style. The relative attitude measure therefore controls for a customer's tendency to report either high or low

[8] In a simulation study, we also found that our model specification and estimation algorithm could satisfactorily recover the true parameters. Further details are available on request.

average attitudes. Weighing both model complexity and inferential benefits, we chose to use average relative attitudes; the existing model is already complex, and our interest is in cross-sectional differences, not dynamics in customer attitudes. Incorporating customer attitude dynamics will mostly likely only improve upon the substantive conclusions of our study but at the cost of a significantly more complex model.

Observed customer heterogeneity can be captured by the number of sales calls from competing firms ($\overline{C}\_Det_i$), the share of wallet of the customer ($\overline{S}OW_i$), and a dummy variable that identifies whether the customer is a specialist (SPC_i). We provide the operationalization of these variables in Table 1. By including the average share of wallet for a customer across time periods in the hierarchical model, we can compare the value of including attitudes, beyond other proxies. Share of wallet is available for all physicians in all time periods due to regulations, though such information is less readily available in other industries and would need to be obtained through customer surveys. Accordingly, we include customer's share of wallet in the hierarchical model along with the covariates that typically would be collected from customer surveys. The hierarchical model for customer i's coefficients at time t are given by

$$\beta_i = \gamma_0 + \gamma_1 \overline{R}ela_i + \gamma_2 \overline{C}\_Det_i + \gamma_3 \overline{S}OW_i + \gamma_4 SPC_i + v_{2it}, \tag{6}$$

where, v_{2it} is a vector of random errors assumed to follow a normal distribution with zero mean and a variance–covariance matrix V_2_i. Furthermore, the γ_1 coefficients capture the effect of relative attitudes on sales and retention: γ_{11} represents the effect of relative customer attitudes on the base sales to the customer (i.e., direct effect); γ_{12}, γ_{13}, γ_{14}, and γ_{15} represent the effects of the interactions of relative customer attitudes with sales calls, recency, frequency, and monetary value on sales; and γ_{16}, γ_{17},....$\gamma_{1,11}$ capture the corresponding influences of the customer's relative attitudes on customer retention. Finally, γ_2, γ_3, and γ_4 take structures similar to γ_1 and capture the influence of competitive sales calls, share of wallet, and customer specialty, respectively.

Endogeneity of sales calls

A firm does not determine the level of sales calls (Det_{it}) in each time period randomly. Instead, it likely allocates sales calls to each customer on the basis of its expectation of the customer's behavior and responsiveness. At a minimum, firms can be expected to pool customers into groups on the basis of their category volume or total sales. This aggregate, group-level sales call allocation guidance then can be updated by individual salespeople, using information they obtain about each customer.

Manchanda et al. (2004) and Donkers et al. (2006) show that the net effect of such allocations could be that sales calls are a function of base-level sales (i.e., the intercept in Equations 3 and 4) and customer responsiveness to sales calls. Ignoring such endogeneity would lead to biased estimates for the response coefficients in the hierarchical Equation 6. We model the level of sales calls directed toward customer i in month t as arising from a Poisson distribution with mean η_{it}[9]:

$$p\left(Det_{it} = m\right) = \frac{\eta_{it}^{m} \exp\left(-\eta_{it}\right)}{m!}, \tag{7}$$

where, m = 0, 1, 2, …, and $\eta_{it} > 0$. We model the natural logarithm of the mean of the Poisson distribution as follows:

$$\ln\left(\eta_{it}\right) = \zeta_0 + \zeta_1 \bar{S}OW_i + \zeta_2 SPC_i + \sum_{k=1}^{11} \zeta_{k+3} \beta_{ki} \tag{8}$$

Share of wallet ($\bar{S}OW_i$) and customer specialty (SPC_i) capture observed customer heterogeneity that may influence the firm's level of sales calls allocated to customer i. The hierarchical coefficients (β_i) capture customer-specific aspects of the firm's sales allocation decisions, unobserved in the data. The coefficient ζ in Equation 8 therefore enables us to infer whether the firm considers customers' sales potential and responsiveness to sales calls in its call plans. The framework proposed here is flexible, without any restrictive assumptions about the process for allocating sales calls. If the firm ignores customer behavior when allocating sales calls, we would not observe a significant effect for any of the ζ coefficients in Equation 8. In addition, we estimated a model with attitudes toward the firm and competition in Equation 8, in addition to the hierarchical equation coefficients. The coefficients of attitudes toward the firm and competition in this modified Equation 8 were not significant, consistent with the pharmaceutical firm's choice not to use customer-level attitude information to determine the level of its sales calls. We therefore do not model sales call per customer as dependent on their attitudes.

We jointly estimate the sales call model (Equations 7 and 8) and the ZIP model of customer retention and sales (Equations 2–6). Because the customer response coefficients (β_{ki})

9 The variance in sales calls was often less than the mean level of sales calls, so that overdispersion was not an issue.

affect the sales model in Equation 8, the posterior distribution of β_{ki}, which is our main focus, is conditional on the distribution of sales model (see Appendix 1). Similar to Donkers et al. (2006), we observe that failing to accommodate the endogeneity of sales calls underestimates the customer response coefficients, β_{ki}, in the ZIP model.[10]

Predicting customer lifetime value

Predicting CLV is challenging in our context, because not all physicians respond to the attitude surveys. We therefore explain our methodology for imputing the attitude values for customers who do not respond to the survey, then explain how to incorporate this imputation into the CLV prediction. To impute missing customer attitudes, we first specify an ordered probit model (Albert and Chib 1993). Let p_ao_{ij} represent the probability that customer i's attitude toward the focal firm (ao_i) equals j, where j = 1, …, 7; that is, $p\_ao_{ij} = P(ao_i = j)$. We define the cumulative probabilities as $\eta_{ij} = \sum_{k=1}^{j} p_{ik}$, which can be modeled as

$$\eta_{ij} = \Phi\left(\omega_j - \delta_{11}\overline{D}et_i - \delta_{12}\bar{y}_i\right), \tag{9}$$

where

ω_j = bin boundaries for each attitude level j,
$\overline{Det}_i$ = average number of sales calls for customer i from t = 1 to T, and
$\bar{y}_i$ = average sales provided by customer i from t = 1 to T.

According to information integration theory, customers combine information from product trials (or prior product experience, $\bar{y}_i$) and marketing communications ($\overline{D}et_i$) to form attitudes (Smith 1993). Cumulative probabilities are necessary until j = 6; the cumulative probability for the seventh category is 1. Let ao_i^* represent a latent continuous random variable, distributed $ao_i^* \sim$ Normal ($x_i^T\delta$, 1). The latent variable can be linked to the observed variables as follows:

$$ao_i = j \text{ if } \omega_{j-1} < ao_i^* < \omega_j \tag{10}$$

We define ω_0 as equal to $-\infty$, and ω_7 is ∞. The regression vector δ and the bin boundaries $\omega_1, \ldots, \omega_6$ are unknown. We impose the restriction $\omega_1 = 0$ to ensure the parameters are identified. A

[10] Further details are available on request.

similar framework can model attitudes toward the competition, with $\delta_2 = (\delta_{2,1}, \ldots, \delta_{2,4})$ as the corresponding coefficients. Differences in the predicted attitudes toward the firm ($\hat{a}o_i$) and competitors ($\hat{a}c_i$) provide the predicted relative customer attitudes ($\hat{r}ela_i$).

We estimate the model parameters in a Bayesian framework employing Markov chain Monte Carlo (MCMC) algorithms to enable posterior inference. The prior specifications for the model parameters and estimation algorithms are in Appendix 1. The posterior predictive distribution of retention and sales for customer i in time period T + k (where k = 1, 2, …, 36) proceeds as follows at the end of every iteration of the MCMC algorithm:

1. Predict relative customer attitudes, $\hat{r}ela_i$, using Equations 9 and 10. Use the predicted relative customer attitudes, ($\hat{r}ela_i$), for customers who did not respond to the other survey, and use the average observed relative customer attitudes for those who did.
2. Predict customer i's hierarchical coefficients $\hat{\beta}_i$ using Equation 6. Relative customer attitudes, $\hat{r}ela_i$, comes from step 1. Competitive sales calls ($\overline{C}\_Det_i$) and share of wallet ($\overline{S}OW_i$) are averages over the calibration time period.
3. Predict the firm's sales calls to customer i in time period T + k, $\hat{D}et_{iT+k}$, using Equations 7 and 8.
4. Predict $\hat{\pi}_{iT+k}$ and $\hat{y}_{iT+K}$ using the predicted coefficients, $\hat{\beta}_i$, and the predicted sales calls, $\hat{D}et_{iT+K}$. $\hat{R}ecency_{iT+K}$, $\hat{F}requency_{iT+K}$, and $\hat{M}onetaryValue_{iT+K}$, predicted in the holdout period, reflect predicted sales in the prior three time periods ($\hat{y}_{iT+K-1}$, $\hat{y}_{iT+K-2}$, $\hat{y}_{iT+K-3}$).

For each iteration of the MCMC algorithm, the predicted values $\hat{\pi}_i = (\hat{\pi}_{T+1}, \hat{\pi}_{T+2} \ldots \hat{\pi}_{T+36})$ and $\hat{y}_i = (\hat{y}_{T+1}, \hat{y}_{T+2} \ldots \hat{y}_{T+36})$ serve to compute the CLV for customer i from Equation 1. The posterior expected CLV for customer i is the Monte Carlo average:

$$E\left[CLV_i\left(\hat{\beta}\_1_i\left(\hat{r}ela_i\right), \hat{\pi}_i, \hat{y}_i\right)\right] = \sum_{l=1}^{np} CLV_i\left(\hat{\beta}\_1_i\left(\hat{r}ela_i\right), \hat{\pi}_i, \hat{y}_i, 1\right) / np \quad (11)$$

where, np refers to the number of posterior iterations.

Results

To obtain our estimation sample, we randomly sampled 1,000 physicians from among the 3,000 who responded at least once for both the own and competitive drugs in our database. The distribution of physician responses to the survey, their level of prescriptions, and the level of sales calls allocated to them was reflective of the wider population.[11] With the MCMC algorithm in Appendix 1, we estimated the coefficients of the proposed model framework, using the estimation sample. Of the 30,000 iterations of the MCMC algorithm, we employed the initial 20,000 as burn-in and the last 10,000 as the posterior sample to make inferences.[12] To assess convergence, we also assessed trace plots and simulated the posterior distribution using five different parallel chains. The multivariate potential scale reduction factor (MPSRF),[13] computed using the posterior sample of five chains that ranged from 1.2 to .9 (across the main effects, moderators, and endogenous customer selection variables), indicated convergence in the posterior sample.

Model comparison

Model Fit. We use the aggregate log conditional predictive ordinate (CPO) and deviance information criterion (DIC) to evaluate the in-sample fit of four comparison models, excluding the naïve model that contained the sample average of sales and retention. The hierarchical Poisson regressions (I and III) included only sales, not retention; the hierarchical ZIP regressions (II and IV) simultaneously modeled sales and retention; and Models I and II excluded attitudes from their hierarchical equations. Model IV therefore represents the proposed hierarchical ZIP model. A higher value of the aggregate log CPO and lower values for DIC would indicate better model fit. In Table 2, our proposed hierarchical ZIP model (Equations 2–9) provides the highest aggregate log CPO (–14,647) and lowest DIC (29,872). Similar to previous research (Horsky et al. 2006), we find that including customer attitudes improves model fit and predictive accuracy. The ZIP framework also provides better in-sample fit than a Poisson regression of sales.

Predictive Accuracy. We assess predictive accuracy with the mean absolute deviation (MAD) between the predicted and observed sales values, as well as the percentage of customers who were correctly classified as either active or dormant (i.e., the hit rate). The reported values are the means of the MAD and hit rates, obtained from each draw of the posterior sample. We compared the predictive capabilities of the models for the in-sample one-period ahead, holdout

[11] We repeated the sampling exercise multiple times and performed the estimation for each random set. The substantive results did not change across samples. The substantive conclusions also were unaffected when we estimated the model with a larger sample.
[12] The iterations lasted about three hours on a standard desktop computer.
[13] Available in the package "BOA" in the R software.

sample one period ahead, and holdout sample twelve periods ahead.

For the in-sample test, we estimated the models using 43 months and predicted sales and retention rates in the 44th month. In Table 2, including attitudes in the hierarchical ZIP model provides better predictions of both sales and retention rates. Next, we randomly sampled 842 customers who were not part of the estimation sample to form the holdout sample, for whom we predicted sales and retention rates in the 44th month to evaluate the one-period ahead MAD and hit rates. Finally, we predicted sales and retention rates from the 32nd to 44th month to evaluate the twelve-period ahead MAD and hit rates. The one-period ahead holdout sample predictions exhibited the same pattern as the in-sample one-period ahead predictions, but the MAD were higher and the hit rates lower for the holdout compared with the in-sample predictions. The hierarchical ZIP model that imputed missing attitudes in the holdout sample also provided the lowest twelve-month ahead MAD (2.1) and the highest twelve-month ahead hit rate (.85). That is, including attitudes in a hierarchical ZIP model improved both the one-step ahead and the twelve-month ahead MAD and hit rates in the holdout sample by .2 and .12, respectively. Knowing customers' attitudes provides better short- and long-term forecasts of customer value.

Influence of customer attitudes

In Table 3, we provide the posterior distributions of the coefficients, γ_1 through γ_6 (Equation 6) that link attitudes to customer responsiveness in the hierarchical ZIP model. The estimated coefficients of the proposed hierarchical ZIP model followed our expectations: Relative customer attitudes exerted a positive and significant influence on retention ($\gamma_{04} = 1.2$, $\alpha < .01$). Customers with higher attitudes toward the focal firm were associated with a higher effect of sales calls on sales ($\gamma_{12} = .02$, $\alpha < .01$) and retention ($\gamma_{14} = .77$, $\alpha < .01$). In addition, the negative effects of recency on sales ($\gamma_{13} = -.04$, $\alpha < .01$) and retention ($\gamma_{18} = -2.9$, $\alpha < .01$) were higher for customers with higher relative attitudes. The effect of monetary value on sales also increased for these customers ($\gamma_{15} = .04$, $\alpha < .01$). Better customer attitudes toward the focal firm, relative to competitors, thus enhance the efficacy of the firm's CRM activities.

Comparing the effect of attitudes on retention and sales

If the effect of attitudes is strong, then coefficients in a model that includes attitudes must differ from those in a model without attitudes. In the top panel of Figure 2, we plot the posterior mean of the coefficient of sales calls on sales (β^{λ}_{1i}) for each customer, obtained from a model that excludes attitudes, compared with the corresponding coefficients from a model that includes those attitudes. The bottom panel of Figure 2 is the corresponding graph for the effect of sales calls on retention (β^{π}_{1i}). This coefficient differs greatly between models that include versus exclude attitudes, more so than the coefficient of sales calls on sales, as indicated by the increase in points away from the 45^{o} line in the bottom panel compared with the top panel of Figure 2. Including attitudes likely has a greater effect on models of customer retention than of sales.

In Table 4 we directly compare the sizes of the effects of customer attitudes on retention and sales. Using the estimates from Table 3, we predict customer response coefficients, $\hat{\beta}_i$, at all combinations of attitudes toward the firm and competition. To control for scale differences in the covariates in Equations 3 and 4, as well as to evaluate the eventual effect of attitudes, we used the coefficients, $\hat{\beta}_i$, to predict the expected probability of customer retention and expected sales and, ultimately, CLV, as specified in Equation 1. We used CLV as a common metric to evaluate the effect size of attitudes on sales and retention, with the assumption that firms eventually must be interested in CLV and whether attitudes exert a CLV effect more through sales or retention.

In the first row of Table 4, we predict the retention rate for the mean values of attitudes toward the firm and competition and of the covariates. We then hold the retention rate constant and predict sales and CLV for all combinations of attitudes toward the firm and competition at one standard deviation above and below the covariates. The first row of Table 4 represents the effect of attitudes on CLV through its effect on sales; the second row holds sales constant and illustrates the effect of attitudes on CLV through retention. The effect size of attitudes derives from the coefficient of variation in predicted CLV values for each combination of attitudes toward the firm and competition—that is, the ratio of the variance in CLV values to the mean CLV values. Unlike variance, it does not increase with higher CLV values.

When sales are constant and retention vary, the coefficient of variation in CLV across all combinations of attitudes toward the firm and competition is 1,211. When retention is constant and sales vary, the coefficient of variation in CLV across all combinations of attitudes toward the firm and competition is 272. From Figure 2 and Table 4, we infer that including information about customer attitudes has a greater effect on customer retention than on sales.

Control factors

The effect of relative attitudes on sales and retention is robust to the inclusion of additional control variables in the hierarchical equation, such as competitor sales calls, share of wallet, and lagged retention probability. Competitor sales calls have a significant negative influence on retention ($\gamma_{36} = -2.6$, $\alpha < .01$) and a significant negative moderating influence on the effect of sales calls on sales ($\gamma_{32} = -.03$, $\alpha < .01$) and retention ($\gamma_{37} = -1.5$, $\alpha < .01$). Competitive sales calls also enhance the negative effect of recency on sales ($\gamma_{33} = -.07$, $\alpha < .01$). When we estimated the model separately for customers for whom the correlation between their recall of the focal firm's sales calls and the actual level of sales call was either higher than .8 or fell between .3 and .5, the substantive results did not change.

Customers with higher share of wallet achieve higher sales ($\gamma_{41} = .57$, $\alpha < .01$) and retention ($\gamma_{36} = 4.4$, $\alpha < .01$). The effect of sales calls on customer retention also increases with customers' share of wallet ($\gamma_{36} = 1.4$, $\alpha < .01$). Regarding behavioral loyalty, the effect of frequency on sales ($\gamma_{34} = .84$, $\alpha < .01$) and retention ($\gamma_{39} = 34.9$, $\alpha < .01$) increases with the share of wallet, and the negative effect of recency on retention diminishes with the share of wallet ($\gamma_{38} = -32.4$, $\alpha < .01$). More important, the estimates in Table 4 show that customer attitudes have a significant influence on CLV, even after we control for share of wallet.

Lagged retention probability has a significant and positive influence on current period retention ($\gamma_{1,13} = 2.6$, $\alpha < .01$), which implies strong state dependence in retention probability. Competitive sales calls reduce the state dependence in retention though ($\gamma_{4,13} = -4.4$, $\alpha < .01$). Similar to Manchanda et al. (2004), we find that specialists provide more prescriptions ($\gamma_{51} = .06$, $\alpha < .01$), though sales calls have a lower effect on retention for specialists ($\gamma_{57} = -2.8$, $\alpha < .01$), who draw more on their prior product experience (i.e., higher coefficient for monetary value).

Endogeneity of Sales Calls. We observe from Table 5 that the firm allocates marketing resources in accordance with physician behavior, such that it devotes more sales calls to physicians who are active ($\zeta_8 = .18$, $\alpha < .01$), exhibit a higher prescription volume ($\zeta_3 = .09$, $\zeta_7 =$

.31, $\alpha < .01$), and offer a higher share of wallet ($\zeta_1 = .65$, $\alpha < .01$) but fewer calls to physicians who are more responsive ($\zeta_4 = -.02$, $\zeta_9 = -.003$, $\alpha < .01$). The firm therefore seems to correct the resources it allocates to each physician according to the response it observes. The significance of some coefficients in the endogeneity of sales call model justifies accounting for the nonrandom nature of the sales calls in Equations 8 and 9 (Manchanda et al. 2004).

Imputing Missing Attitudes. The results in Table 6 indicate that, as expected, customers' prior sales (or product experience) and the firm's sales calls have positive and significant effects on customer attitudes toward the firm ($\delta_{11} = 1.18$, $\alpha < .01$; $\delta_{12} = .05$, $\alpha < .01$, respectively) but a negative, significant effect on the attitudes toward the competition ($\delta_{21} = -1.09$, $\alpha < .01$; $\delta_{12} = -.13$, $\alpha < .01$, respectively).

The estimates in Table 6 and the holdout predictions in Table 2 thus reveal that a model including only prior sales and sales calls provides reasonable imputations of missing customer attitudes. These variables are readily available in CRM databases and do not require additional data collection efforts. Other factors, such as competitive sales calls and the quality of the firm's marketing actions, can affect customer attitudes, but the holdout predictions show that our model improves on a model that excludes customer attitudes (Models I and II).

Robustness of results

Sampling Error in Attitude Responses. Confidence in the measures of customer attitude increases if the responses come from multiple periods and the same customer. Survey responses in later periods may be more predictive of their future attitudes than their initial responses; such recent assessments of customer attitudes are especially important for predicting missing customer attitudes in the holdout sample. Therefore, we expect confidence in customer attitude measures to vary with the number and recency of customer responses to attitude surveys.

To assess the sensitivity of our results to sampling errors, we computed a weighted measure of relative attitudes, equal to the difference between weighted attitudes toward the own versus competitors' drugs. Weighted customer attitudes for a drug are more important to customers who have responded recently and more frequently across various surveys, as we detail in Appendix 2. We then inserted the weighted relative attitudes in Equation 6 and obtained results similar to those in the proposed model.

To assess the sensitivity of our predictive results to sampling error, we also computed weighted relative attitudes in the holdout sample. Weighted attitudes toward the own drug among customers in the holdout sample who responded at least once to the surveys were derived in a way similar to the computation in the estimation sample (for the predictions for customers who did not respond at all, see Appendix B). Similar to the estimation sample, the one- and twelve-period ahead predictions in the holdout sample with weighted relative attitudes were similar to a holdout sample with relative attitudes. Therefore, sampling error is not a major concern in our substantive conclusions.

Physician Learning. Customers' uncertainty about a drug's performance can decrease over time with more usage, leading to higher customer attitudes and higher sales call effectiveness. If attitudes are driven by customer learning, they should not have an effect (or at least have a diminished effect) in a model that accommodates customer learning. We estimated a dynamic hierarchical model of physician prescriptions by allowing the coefficients in Equations 3 and 4 to vary over time and by customer.[14]

Relative customer attitudes exerted a positive effect on the effectiveness of detailing for sales and retention even in the dynamic model. The dynamic coefficients did not indicate any systematic time trends, perhaps because we collected our data about a year after the launch of the own drug. Without any new drug launches in the therapeutic category during our study period, physician learning likely was not a major factor in our study. The significant effects of relative customer attitudes even in the dynamic model also show that physician learning cannot account for the effect of attitudes on sales and retention. Because our goal is to establish the value of cross-sectional differences in relative customer attitudes, we use the simpler static framework.

Customer Selection And Resource Allocation

The ongoing collection of customer attitudes demands substantial investments; our focal firm invests more than $1 million annually in such efforts. These investments are justified only if the financial returns of using customer attitudes in customer-level models exceed the investments required to collect the information. To evaluate the value customer attitude information, we use

[14] Details on the dynamic model are available on request.

customer selection and *customer-level resource allocation* (Reinartz and Venkatesan 2008). Specifically, we assess whether different customer responsiveness estimates obtained from a model that includes customer attitudes (Figure 2) also translate into an identification of more profitable customers or a more appropriate allocation of resources to customers.

For both the customer selection and resource allocation exercises, we used the holdout sample of 842 customers from our model comparisons and predicted sales and retention for months 10–45, with the same methods we previously used to compute the customer's CLV in Equation 1. We imputed any missing attitudes in the holdout sample using the model in Equations 10 and 11 and the estimates in Table 6. From conversations with the focal firm and secondary research, we assume a gross margin for the own drug of $120 (equivalent to 70% of the retail price) and a $50 unit cost of a sales call. In Table 7 we report the results from our selection exercise.

Customer selection

The objective of the customer selection process is to identify customers who might be profitable in the future to prioritize them for targeting. In the first column of Table 7, we classify customers into quintiles, according to their observed CLV. Then in the second column, we classify quintiles on the basis of the predicted CLV from the hierarchical ZIP model that includes customer attitudes (i.e. $\hat{C}LV_{ia}$). Finally, in the third column of Table 7, customers are classified into quintiles based on the predicted CLV from a hierarchical ZIP model that did not include customer attitudes, $\hat{C}LV_{ea}$. The total observed CLV for customers identified to be in the top quintile by $\hat{C}LV_{ea}$, is $972,053. The difference between $\hat{C}LV_{ia}$ and $\hat{C}LV_{ea}$ in the fourth column of Table 7, provides an estimate of the incremental profits that the firm would obtain if it selected customers for sales calls using a model that included customer attitudes.

Specifically, the customers in the top quintile by $\hat{C}LV_{ia}$ provide $9,611 more over three years than customers identified to be in the top quintile by $\hat{C}LV_{ea}$. This incremental profit is equivalent to .93% of the total CLV obtained from top quintile customers in terms of their observed profits. That is, if the firm targeted only the top quintile of its customer base, the returns it would obtain by including customer attitudes in CLV models would be .93% higher than if it ignored customer attitudes. Similarly, the returns from including customer attitudes would increase by 3.57%, 29.62%, 79.33%, and 24.12% if the firm targeted only the second, third, fourth, and the final quintile, respectively. Customer attitudes are valuable for customer selection only if the firm targets a subset of customers with sales calls though. We repeated the selection exercise by classifying customers into deciles and find similar results.[15]

Customer resource allocation

The second CRM strategy we evaluated considered the optimal sales force allocation to maximize the CLV of each customer. For each customer, we fixed the estimated coefficients that we used to predict CLV in the holdout sample and varied the sales calls to find the optimal average level of sales calls per month ($\overline{\mathrm{D}}\mathrm{et}_i^*$) that would maximize respective the CLV ($\hat{\mathrm{C}}\mathrm{LV}_i^*$). Our optimization method therefore uses each customer's responsiveness (estimated coefficients) and profit potential (objective function in Equation 1) to allocate resources.[16] For simplicity, sales calls per month remained constant over 36 months. We also restricted the total budget available for the sales call to equal the observed sales call investments over the 36 months in the holdout sample. Therefore, we can compare how different models would reallocate resources with similar resource constraints. Our optimization algorithm allocated sales calls to individual customers with the objective of maximizing the sum of CLVs across all customers, keeping the total budget available for sales calls the same over time.

We used the Nelder-Mead hill-climbing algorithm to optimize sales calls in each scenario. Similar to the selection exercise, we evaluated optimal sales calls and maximized CLV for the hierarchical ZIP model that customer attitudes ($\overline{\mathrm{D}}\mathrm{et}_{i,ia}^*$, $\hat{\mathrm{C}}\mathrm{LV}_{i,ia}^*$ respectively) and the hierarchical ZIP model that did not include customer attitudes ($\overline{\mathrm{D}}\mathrm{et}_{i,nia}^*$, $\hat{\mathrm{C}}\mathrm{LV}_{i,nia}^*$ respectively). The optimal CLV with and without attitudes,($\hat{\mathrm{C}}\mathrm{LV}_{i,ia}^*$, and $\hat{\mathrm{C}}\mathrm{LV}_{i,nia}^*$) are based on estimates from the hierarchical ZIP model that include attitudes and the hierarchical ZIP model without attitudes respectively.

[15] We repeated this exercise with multiple, randomly selected holdout samples, and the conclusions were unaffected.

[16] The correlation between customer responsiveness and the optimal level of detailing was approximately .7; our optimization algorithm thus considered customer responsiveness when allocating resources. Further details are available on request.

We provide the observed sales calls, CLV, optimal sales calls, and maximized CLV for the hierarchical ZIP with and without attitudes in Table 8. We summarize the resource allocation results into quintiles, to clarify the strategic implications of the optimal resource allocation process. Whereas the first column in Table 8 represents quintiles classified by observed CLV, the other columns provide the total sales calls investment per customer in each quintile and the total CLV of customers in each quintile. As we expected, the firm allocates the most sales calls to customers with the highest CLV (36 total sales calls in 36 months) and decreases its sales call investment when the quintiles increase from 2 to 5. The optimal sales call investment in the hierarchical model without customer attitudes follows a pattern similar to the firm's current strategy. However, the hierarchical ZIP model with attitudes recommends allocating more resources to customers in the middle tiers (quintiles 3 and 4). The uncertainty in the model predictions, indicated by the 90% confidence interval for the maximized CLV from the hierarchical model, is not too high.

Different customer profitability tiers

The selection and resource allocation exercises reveal that customer attitudes provide more diagnostic value for mid-tier customers than for others. The incremental profits (maximized CLV) obtained by including customer attitudes are highest for these customers,[17] so firms seemingly should allocate more resources to them. The variation in attitudes toward the firm and competitors across these quintiles may help explain the results in Tables 7 and 8.

The most valuable tier has the highest attitudes toward the firm (average 5.7) and lowest attitudes toward the competition (5.4). These customers exhibit comparably less variance (.61 for the firm, .59 for the competition)—maybe because of the strong fit between customer needs and the focal firm's offering. The least valuable tier instead exhibits the lowest attitudes toward the firm (5.2), and highest attitudes toward the competition (5.6). Again we observe little variance within customers in this group (.69 for the firm, .72 for the competition). The mid-tier customers have moderate attitudes regarding both the firm (5.4) and the competition (5.5), but the variance in their attitudes are highest of all groups (.85, .83). Therefore, we argue that the heterogeneity of this group's attitudes also is greater than in the other two groups, and for the firm, it pays most to exploit such heterogeneity.

[17] We repeated this exercise with multiple, randomly selected holdout samples, and the conclusions were unaffected.

In the top and bottom tiers, attitudes align with behavior, so knowledge of attitudes provides relatively few benefits; most of the information already is encapsulated in behavior. Therefore, prior transactions with the firm are sufficient indicators of future value. In contrast, mid-tier customers differ greatly in their attitudes, and knowledge of these attitudes yield relatively higher benefits in terms of managing customers (i.e., selecting for targeting).

The results in Table 8 related to the optimization algorithm also suggest that for top-tier customers, it is better to reduce sales calls from current levels (and thus reduce costs) and simply capitalize on their higher inherent brand preference and positive prior product experiences (i.e., higher past sales, lower recency, and higher frequency) to improve profits. Middle-tier customers have moderate intercepts and coefficients of sales calls, in line with their moderate levels and higher variance in attitudes toward both the firm and competitors. The algorithm therefore suggests that the firm should invest in more sales calls among customers with higher attitudes in the middle tier, because brand preferences and positive prior experience have a weaker effect. Finally, the lower-tier customers exhibit the lowest intercept and coefficient of sales calls, but their sales and retention levels are not high enough to justify more sales calls per month. In summary, when facing resource constraints, managers should focus their efforts on collecting the attitudes of their mid-tier customers, rather than others.

Validation of resource allocation strategy

To validate this proposed resource allocation strategy, we followed Zhang and Krishnamurthi (2004) and Khan et al. (2009). We classified customers in the holdout sample into match, missed opportunities, and wasted resources categories, according to the difference between the modeled optimal monthly sales calls (including or excluding attitudes) and the observed average monthly level of sales calls allocated by the firm during the first 12 months (10–22) of the prediction window.[18] Customers who fell within the 40th to 60th percentiles were categorized as matched; their optimal and observed sales calls were similar (with a difference close to 0). Customers below the 40th percentile represented wasted sales calls; customers above the 60th percentile entailed missed opportunities, because more sales calls would be optimal.

[18] We do not use CLV to validate the optimization strategy, because the firm reviews its sales calls allocations every year, and three years is a long time to assess the validity of the optimal strategy. Although the allocation uses a long-term, customer value perspective, reviews are generally done on an annual basis.

We report the number of customers and average 12-month profit (month 10–22) for each segment in Table 9. In the model that includes attitudes, the average profits are highest ($56) for the 258 customers classified in the match segment class, followed by the 190 customers in the missed opportunities ($51) and the 366 customers in the wasted sales calls ($43) segments. In contrast, the model that excludes attitudes indicates that average profits are highest among the 197 missed opportunities customers ($54), followed by the 245 match segment customers ($52) and then the 372 wasted sales calls customers ($42). We thus validate our resource allocation recommendations, because the highest profits among all groups come from customers for whom the actual firm practice matched the recommendations of a model that includes attitudes.

Managerial Implications

If companies include information about attitudes in their behavioral CLV models, they can both explain and predict customer responses better (sales, retention, CLV). The positive effect of better attitudes on CLV works predominantly through the retention aspect. This result has important implications for the kind of metrics and performance indicators that firms use to diagnose their customer base. For example, firms should consider customer retention rates rather than sales for any campaign focused on improving the customer's emotional attachment. Our results indicate that a diagnosis on the basis of sales is comparatively less useful.

A key result from the selection and resource allocation exercises states that the incremental profits from including customer attitudes are greatest for mid-tier customers. Yet in existing CRM literature, this segment is probably the least discussed group of customers. Including information on attitudes in CLV models is diagnostic for this group and yields significant benefits, likely because these customers' attitudes provide a forward-looking measure that effectively can discriminate mid-tier customers with the potential to grow from those whose profitability is likely to decrease. Our results also imply that firms may be overspending on top-tier customers. When customer attitude information is available, firms can improve the ROI of their CRM campaigns by balancing resources between top- and mid-tier customers. This new finding has not previously emerged in CRM literature; it extends an existing perspective that states resources should be devoted primarily to top customers (Rust et al. 2004).

Finally, we demonstrate the positive returns from including attitudinal information into CLV modeling. It depends on the incremental lift that the model provides but also acknowledges the cost of collecting additional data. In our case, these returns are substantial, especially for mid-tier customers. We can project the incremental return estimate from the holdout sample to the entire customer base of more than 150,000 customers to obtain the financial ROI for the data collection. That is, using customer attitudes to select the top three quintiles across the entire customer base, we calculate an incremental annual profit of $32 per customer, equivalent to a projected annual return of more than 300% on the $1 million investment for collecting customer attitudes. The strong customer selection capability thus establishes that the benefits from using customer attitudes more than pay for the investments required to collect this information.

Limitations and Further Research

An understanding of customer attitude formation that includes the effects of mass marketing and the importance of the marketing message would provide a good contribution. Examining the interaction between marketing activities intended to build customers' emotional attachment to a brand and CRM activities geared toward developing a long-term, profitable relationship also represents a promising avenue. It is possible that the benefits provided by customer attitudes exhibit diminishing returns; therefore, further research might assess the dynamic interplay of marketing tactics that improve customer attitudes versus CRM activities that adopt a dynamic optimization procedure. Evaluating such threshold effects could help firms optimize the balance between investing to develop customer attitudes and investing in customer retention. Additional research should investigate the utility of including an explicit component to model the probability that a customer responds to an attitude survey, then assess the potential bias related to the impact of attitudes. We evaluate the impact of customer attitudes on customer sales, which is just one element of customer profitability and CLV. Additional research should also examine the influence of customer attitudes on share of wallet and word-of-mouth behavior.

REFERENCES

Ahearne, M., R. Jelinek, and E. Jones (2007), "Examining the Effect of Salesperson Service Behavior in a Competitive Context," *Journal of Academy of Marketing Science*, 35(4), 603-616.

Alba, J. W., J. Wesley, and J. Lynch Jr. (1991), "Memory and Decision Making," in *Handbook of Customer Behavior*, T. Robertson and H. Kassarjian, eds. Englewood Cliffs, NJ: Prentice Hall.

Albert, J. H., and S. Chib (1993), "Bayesian Analysis of Binary and Polychotomous Response Data," *Journal of American Statistical Association*, 88(422), 669-679.

Bijmolt, T. H.A., P.S.H. Leeflang, F. Block, M. Eisenbeiss, B.G.S. Hardie, A. Lemmens, and P. Saffert (2010), "Analytics for Customer Engagement," *Journal of Service Research*, 13(3), 341-356.

Donkers, B., R. Pacep, J. Jonker, and P.H. Frances (2006), "Deriving Target Selection Rules From Endogenously Selected Samples," *Journal of Applied Econometrics,* 21(5), 549-562.

Drèze, X. and J.C. Nunes (2009), "Feeling Superior: The Impact of Loyalty Program Structures on Consumer's Perceptions of Status," *Journal of Consumer Research*, 35 (6), 890-905.

Gatignon, H. and J.-M. Xuereb (1997), "Strategic Orientation of the Firm and New Product Performance," *Journal of Marketing Research*, 34-1, 77-90

Ghosh, S.K., P. Mukhopadhyay, and J.C. Lu (2006), "Bayesian Analysis of Zero-Inflated Regression Models," *Journal of Statistical Planning and Inference*, 136(4), 1360-1375.

Gupta, S., D.R. Lehmann and J.A. Stuart (2004), "Valuing Customers," *Journal of Marketing Research*, 41(1), 7-18.

Gupta, S. and V. Zeithaml (2006), "Customer Metrics and Their Impact on Financial Performance," *Marketing Science*, 25(6) 718–739.

Horsky, D., S. Misra, and P. Nelson (2006), "Observed and Unobserved Preference Heterogeneity in Brand-Choice Models," *Marketing Science*, 25(4), p. 322-335.

Kerin, R. A. and R. O. Regan (2008), *Marketing Mix Decisions: New Perspectives and Practices*. Chicago: American Marketing Association.

Khan, R., M. Lewis, and V. Singh (2009), "Dynamic Customer Management and the Value of One-to-One Marketing," *Marketing Science*, 28(6), 1063-1079.

Lambert, Diane (1992), "Zero-Inflated Poisson Regression, with an Application to Defects in Manufacturing," *Technometrics*, 34(1), 1-14.

Manchanda, P., P.E. Rossi, and P. Chintagunta (2004), "Response Modeling with Nonrandom Marketing-Mix Variables," *Journal of Marketing Research*, 41(November), 467-478.

Mizik, N. and R.L. Jacobson (2009), "The Financial Markets Research in Marketing," *Journal of Marketing Research*, 46 (3), 320-324.

Netzer, O., J.M. Lattin, and V. Srinivasan (2008), "A Hidden Markov Model of Customer Relationship Dynamics," *Marketing Science*, 27(2), 185-204.

Petty, R.E., R. Unnava, and A.J. Strathman (1991), "Theories of Attitude Change," in *Handbook of Customer Behavior*, T. Robertson and H. Kassarjian, eds. Englewood Cliffs, NJ: Prentice Hall.

Reinartz, W. and R. Venkatesan (2008), "Decision Models for Customer Relationship Management (CRM)," in *Handbook of Marketing Decision Models*, Berend Wierenga (ed.), New York: Springer, 291-326.

Rust, R., K.N. Lemon, and V.A. Zeithaml (2004), "Return on Marketing: Using Customer Equity to Focus on Marketing Strategy," *Journal of Marketing*, 68(1), 109-127.

Rust, R. and P. Verhoef (2005), "Optimizing the Market Interventions Mix in Intermediate Term CRM," *Marketing Science*, 24(3), 477-489.

Schmittlein, D. C. and D.G. Morrison (1985), "Is the Customer Still Active?" *American Statistician*, 39(4), 291-295.

Smith, R.E. (1993), "Integrating Information From Advertising and Trial: Processes and Effects on Consumer Response to Product Information," *Journal of Marketing Research*, 30(May), 204-19.

Srinivasan, S. and D. Hanssens (2009), "Marketing and Firm Value: Methods, Metrics, Findings, and New Directions," *Journal of Marketing Research*, 46(3), 293-312.

Srinivasan, S., M. Vanhuele, and K. Pauwels, (2010) "Mind-Set Metrics in Market Response Models: An Integrative Approach," *Journal of Marketing Research*, 47(4), 672-84

Stahl, F., M. Heitmann, D.R. Lehmann, and S.A. Neslin (2010), "The Impact of Brand Equity on Customer Acquisition, Retention, and Profit Margin," MSI Working Paper Series, 10-116.

Venkatesan, R. and V. Kumar (2004), "A Customer Lifetime Value Framework for Customer Selection and Resource Allocation Strategy," *Journal of Marketing*, 68(4), 106-125.

Zhang, J., and L. Krishnamurthi (2004), "Customizing Promotions in Online Stores," *Marketing Science*, 23(4), 561-578.

TABLE 1

VARIABLE OPERATIONALIZATION AND DESCRIPTIVE STATISTICS

Variable	Operationalization	Mean	Standard Deviation
$Sales_{it}$	Number of new prescriptions of the firm's drug provided by customer i in month t	3.1	4.6
Det_{it}	Number of sales calls from the firm directed to customer i in month t	2.1	1.8
$Recency_{it}$	Time since customer i's last prescription of the firm's drug in the three months prior to t	1.5	.99
$Frequency_{it}$	Number of instances with at least one prescription of the firm's drug by customer i in the three months prior to t	2.1	1.03
Monetary $Value_{it}$	Total number of prescriptions of the firm's drug by customer i in the three months prior to t	7.7	4.1
$P(Alive)_{it-1}$	Probability that customer i will actively prescribe the firm's drug after t-1	.87	.25
$\overline{Rela}_i$	Difference between the average attitude of customer i toward the firm and the average attitude of customer i toward the competition, with the average calculated over customer i's survey responses during the data collection time frame, from t = 1 to T.	-.47	1.1
$\overline{C\_Det}_i$ [a]	Average of customer i's estimates of the number of sales calls from competing firms, calculated over customer i's survey responses during the data collection time frame, from t = 1 to T.	2.2	2.8
$\overline{SOW}_i$	Average of the firm's share of wallet of customer i obtained as the ratio of the total sales of the firm's drug to the total category sales, calculated over the data collection time frame, from t = 1 to T.	.18	.21
SPC_i	Equal to 1 if the customer is a specialist, and 0 otherwise.	.45	.55

[a] Customer estimates obtained from the survey and average of sales calls of competition.

TABLE 2

MODEL COMPARISON

	DIC	Aggregate Log CPO	PSBF*	Sales			Retention		
				In-Sample	Holdout Sample		In-Sample	Holdout Sample	
				MAD* - One Period Ahead	MAD - One Period Ahead	MAD - Twelve Months Ahead	HR - One Period Ahead (%)	HR - One Period Ahead (%)	HR - Twelve Months Ahead (%)
Sample Average				3.1	3.7	3.54	57	49	49
No Attitudes in Estimation and Prediction									
Hierarchical Poisson	38,271	-19,152	4,505	2.6	3.3	3.6	NA		
Hierarchical ZIP	30,729	-15,257	610	1.7	1.9	2.3	83	75	73
Include Attitudes for Estimation and Predict Attitudes in Holdout Sample									
Hierarchical Poisson	38,067	-18,615	3,968	2.3	2.7	3.3	NA		
Hierarchical ZIP	29,872	-14,647	NA	1.4	1.7	2.1	91	87	85

Notes: DIC = deviance information criterion; CPO = conditional predictive ordinate; PSBF = pseudo Bayes factor; MAD = mean absolute deviation; HR = hit rates; ZIP = zero-inflated Poisson model. If the PSBF is greater than 20, the model with the larger aggregate log CPO provides better fit. Difference in DIC values between two models that are greater than 10 support the model with the lower DIC value. NA = Not applicable.

** Reported MAD values are the average of the MAD obtained from each draw of the posterior sample.

TABLE 3

ESTIMATES OF THE HIERARCHICAL MODEL

On :	Sales					Retention					
Effect of:	**Intercept**	**Sales Calls**	**Recency**	**Frequency**	**Monetary Value**	**Intercept**	**Sales Calls**	**Recency**	**Frequency**	**Monetary Value**	**Lagged Retention Probability**
Intercept	.33 (.08)	.11 (.01)	-.02 (.008)	.13 (.01)	.03 (.0003)	n.s.	28.8 (2.4)	-.19 (.02)	.24 (.02)	.06 (.003)	.71 (.06)
Relative Firm Attitudes	n.s.	.02 (.002)	-.04 (.01)	n.s.	.04 (.02)	1.2 (.9)	.77 (.2)	-2.9 (1.7)	n.s.	n.s.	n.s.
Competitive Sales Calls	n.s.	-.03 (.004)	-.007 (.002)	n.s.	n.s.	-2.6 (.99)	-1.5 (.3)	n.s.	n.s.	n.s.	-4.4 (2.7)
Share of Wallet	.57 (.05)	n.s.	n.s.	.84 (.55)	n.s.	4.4 (1)	1.4 (.33)	-32.4 (27.1)	34.9 (22.6)	n.s.	n.s.
Specialty	-.06 (.01)	n.s.	n.s.	n.s.	n.s.	n.s.	-2.8 (.36)	n.s.	n.s.	-3.3 (1.2)	n.s.

Notes: Values in parentheses are posterior standard deviations. n.s. = not significant. The estimates are considered significant if at least 95% of the posterior does not contain 0.

TABLE 4

EFFECT OF ATTITUDES ON RETENTION AND SALES

	Coefficient of Variation in CLV for all Combinations of Firm and Competitive Attitudes	
	Covariates One Standard Deviation Above Mean	**Covariates One Standard Deviation Below Mean**
Sales Varying, Retention Constant	386.16	159.62
Retention Varying, Sales Constant	674.51	1749.31

TABLE 5

ENDOGENEITY OF SALES CALLS

	Mean (standard deviation)
ζ_0	1.3 (.06)
ζ_1	.65 (.11)
ζ_2	n.s.
ζ_3	.09 (.01)
ζ_4	-.02 (.003)
ζ_5	.42 (.11)
ζ_6	n.s.
ζ_7	.31 (.05)
ζ_8	.18 (.06)
ζ_9	-.003 (.0001)
ζ_{10}	n.s.
ζ_{11}	n.s.
ζ_{12}	-.05 (.002)

Notes: n.s. = not significant. The estimates are considered significant if at least 95% of the posterior does not contain zero.

TABLE 6

IMPUTING MISSING CUSTOMER ATTITUDES

	Firm Attitudes	Competitive Attitudes
Intercept	1.34 (.35)	3.72 (.296)
Lagged Firm Sales	1.18 (.2)	-1.09 (.18)
Lagged Firm Sales Calls	.05 (.01)	-.13 (.03)

Notes: Values in parentheses are posterior standard deviations. All the estimates are significant, because at least 95% of the posterior does not contain 0.

TABLE 7

VALUE OF KNOWING CUSTOMER ATTITUDES IN CUSTOMER SELECTION

Customer Quintile	Total Observed CLV: Customers Rank Ordered Based on Observed CLV	Total Observed CLV: Customers Rank Ordered Based on Predicted CLV-from *Model Including Attitudes*	Total Observed CLV: Customers Rank Ordered Based on Predicted CLV from *Model Excluding Attitudes*	Incremental Profit from Including Attitudes	Percentage Lift in Profit from Using Attitudes
	[1]	*[2]*	*[3]*	*[4]=[2]-[3]*	*[5]=[4]/[1]*
1	1,038,515	981,664	972,053	9,611	.93%
2	333,586	313,999	302,081	11,918	3.57%
3	119,107	170,592	135,316	35,276	29.62%
4	-25,198	-22,247	-2,258	-19,989	79.33%
5	-152,666	-130,663	-93,847	-36,816	24.12%

TABLE 8

VALUE OF KNOWING CUSTOMER ATTITUDES IN CUSTOMER-LEVEL RESOURCE ALLOCATION

Customer Quintile	Observed Total Sales Calls per Customer	Optimal Total Sales Calls per Customer without Using Attitudes	Optimal Total Sales Calls per Customer Using Attitudes	Observed CLV	Predicted CLV* without Attitudes	Predicted CLV* with Attitudes	10th Percentile Predicted CLV* with Attitudes	90th Percentile Predicted CLV* with Attitudes
1	36	33	26	1,038,515	2,817,108	3,291,260	2,929,222	3,653,300
2	29	25	25	333,586	904,532	1,724,386	1,534,703	1,914,069
3	24	23	30	119,107	347,255	964,907	858,768	1,071,047
4	19	22	28	-25,198	101,351	325,717	289,889	361,546
5	14	21	23	-152,666	52,636	94,870	84,435	105,306

Notes: CLV* indicates maximized customer lifetime value.

TABLE 9

VALIDATION OF RESOURSE ALLOCATION STRATEGY INFORMED BY CUSTOMER ATTITUDES

	Resource Allocation Model			
	Including attitudes		Excluding attitudes	
	Number of Customers	12-Month Profit	Number of Customers	12-Month Profit
Match	258	**56**	245	52
Missed Opportunities	190	51	197	**54**
Wasted Sales Calls	366	43	372	42

FIGURE 1

CONCEPTUAL MODEL OF THE IMPACT OF ATTITUDES

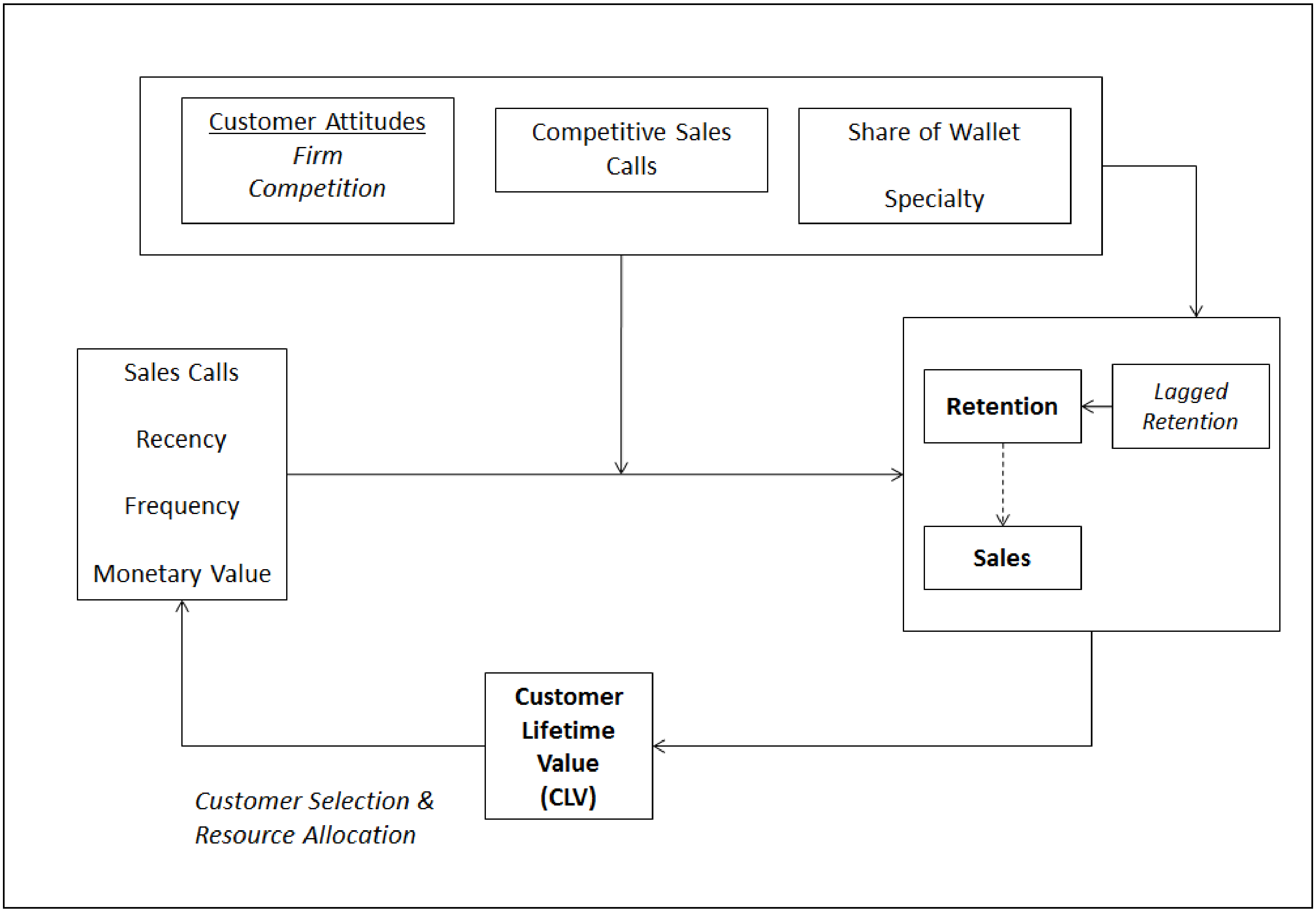
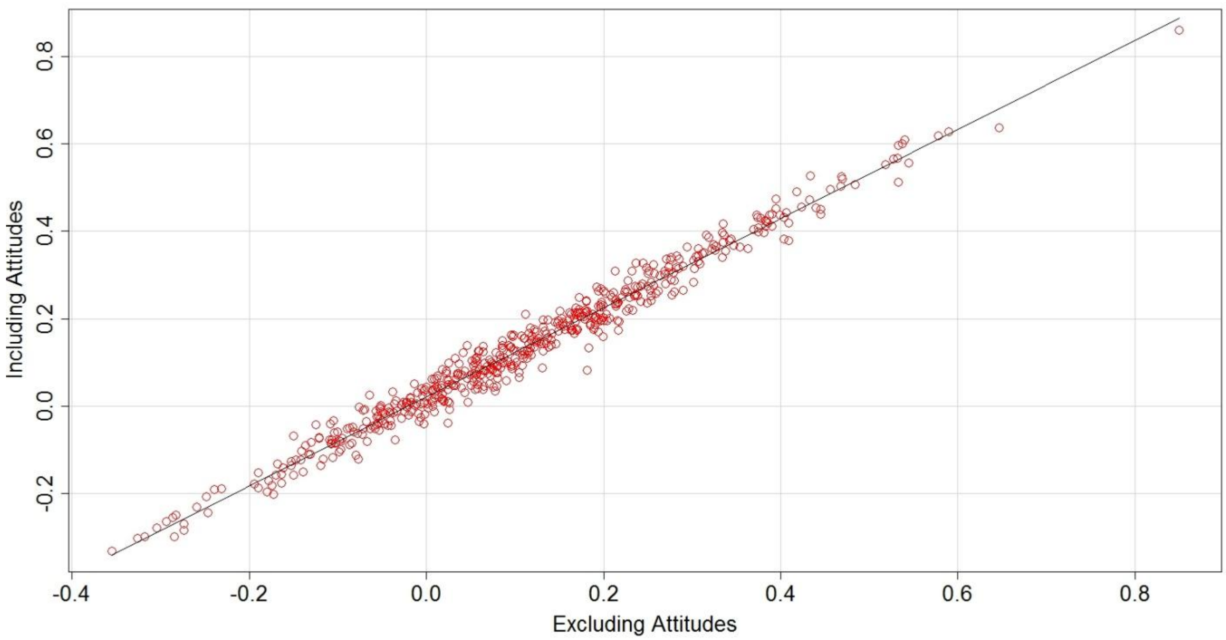
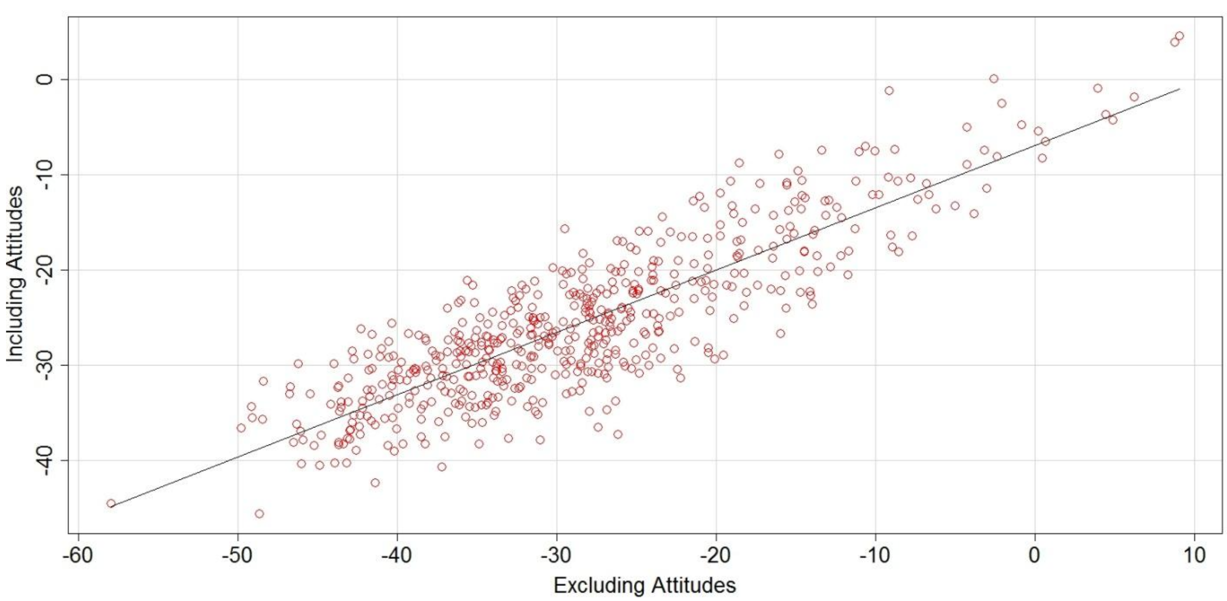


FIGURE 2
COMPARISON OF SALES CALL EFFECTS
Effect of Sales Calls on Sales
(β_{1i}^{λ})



Effect of Sales Calls on Retention

(β_{1i}^{π})



APPENDIX 1

CONDITIONAL DISTRIBUTIONS AND MCMC ESTIMATION ALGORITHMS

The likelihood for our model framework (Equations 1–7) can be determined according to the following equation:

$$L\left(Y^T \mid \beta\_1_i, \beta\_2_t, \gamma, \varsigma, \alpha\right) = \prod_{t=1}^{T}\prod_{i=1}^{N} \begin{matrix} p_{ZIP}\left(Y_{it} \mid \lambda_i, \pi_i, X_i^{\pi}, X_i^{\lambda}, \beta_i\right) p_{poisson}\left(Det_{it} \mid \varsigma, \beta_i\right) \\ p_{normal}\left(\beta_i \mid \overline{Rel}a_i, \overline{SOW}_i, \bar{C}\_Det_i, SPC_i, \gamma\right) \end{matrix}.$$

Let Y_t denote an $N \times 1$ vector of new prescriptions observed across all customers at time t; $Y_t = (y_{1t}, y_{2t}, \ldots, y_{Nt})'$, where $'$ denotes a transpose. Similarly, let $\lambda_i = (\lambda_{i1}', \lambda_{i2}', \ldots, \lambda'_{iT})'$, $\pi_i = (\pi_{i1}', \pi_{i2}', \ldots, \pi'_{iT})'$, and $Det_i = (Det'_{i1}, Det'_{i2}, \ldots, Det'_{iT})'$, each denote a $T \times 1$ vector of the mean of the Poisson distribution of new prescriptions, mean of the Bernoulli distribution for customer retention, and the level of sales calls allocated for customer i in all time periods respectively. Then X_{it}^{λ} and X_{it}^{π} denote the $1 \times P^{\lambda}$ and $1 \times P^{\pi}$ row vectors of covariates in the regression functions of $\ln(\lambda_{it})$ and $\text{logit}(\pi_{it})$, as represented in Equations 3 and 4.

Prior Specifications

Let D_0 denote all the information available at time $t = 0$. We specify fairly standard, conditionally conjugate prior distributions. We assume an inverted Wishart prior for the error variance in the hierarchical model, $V_\beta \sim IW(\text{diag}(10), (p^{\lambda}+p^{\pi})+3)$. Given p^{λ} and p^{π} are 5 and 6, respectively, the prior specifications for V_β is $V_\beta \sim IW(\text{diag}(10), 14I_{11})$. We assume multivariate normal distributions as priors for γ and ζ. The prior distributions are specified as $\gamma \sim MVN(0, \text{diag}(1000))$ and $\zeta \sim MVN(0, \text{diag}(1000))$.

Data Augmentation

We use the latent variable structure outlined in Equation 2b for model estimation. Instead of obtaining samples directly from the posterior of (π, λ) given data Y, the sampling instead comes from the posterior of (π, λ, V, B) given data Y. To implement this procedure within a Gibbs sampling algorithm, the conditional distribution of (V, B) given (Y, π, λ) is required; sampling from this conditional distribution is commonly referred to as "data augmentation." Ghosh et al. (2006) show that

$$
(V = k, B = 0 \mid y) = \begin{cases} \dfrac{(1-\pi)p(V=0)}{\pi + (1-\pi)p(V=0)} & if \ \mathrm{v = y = 0} \\ 1 & if \ \mathrm{v = y > 0} \\ 0 & otherwise \end{cases}
$$

and,

$$
(V = k, B = 1 \mid y) = \begin{cases} \dfrac{\pi p(V=k)}{\pi + (1-\pi)p(V=0)} & if \ \mathrm{y = 0} \\ 0 & otherwise \end{cases}
$$

If $y_{it} = 0$ is observed, we can draw a random B_{it} from a Bernoulli distribution with probability π, equivalent to flipping a coin with probability of HEAD being π. If the random draw $B_{it} = 1$, then we draw V_{it} from a Poisson distribution with parameter λ_{it}. If the random draw $B_{it} = 0$, then we set $V_{it} = 0$. If $y_{it} > 0$, $B_{it} = 0$ and $V_{it} = y_{it}$.

Conditional Distributions

The Gibbs sampler proceeds by sequentially sampling from conditional distributions of the parameters.

Sampling from the Conditional Distribution of β_i. Because we assume a Poisson distribution for sales or new prescriptions and a Bernoulli distribution for retention, we use Metropolis-Hastings algorithm to sample β_i. Let β_i^p denote the previous draw for β_i. For each customer i, a candidate value β_i^c is sampled from a normal distribution provided by $N(0,k\Psi)$. We set Ψ to the asymptotic variance–covariance matrix of the b parameters estimates on pooled data, similar to Manachanda et al. (2004). Let Z_i denote the set of covariates in Equation 6. We extract the candidates for sales and retention, $\beta^{\lambda c}{}_i$ and $\beta^{\pi c}{}_i$, from $\beta^c{}_i$. The acceptance probability for $\beta^{\lambda c}{}_i$ is given by

$$
p(accept) = \min\left(\frac{p_{Poisson}(V_{it} \mid \lambda_{it}^c) * p_{normal}(\beta_i^c \mid Z_i, \gamma) * p_{poisson}(Det_{it} \mid \eta_{i,t}^c)}{p_{Poisson}(V_{it} \mid \lambda_{it}^p) * p_{normal}(\beta_i^p \mid Z_i, \gamma) * p_{poisson}(Det_{it} \mid \eta_{i,t}^p)}, 1 \right),
$$

where λ_i^c and λ_i^p are obtained from Equation 3 using $\beta_i^{\lambda c}$ and $\beta_i^{\lambda p}$, respectively.

The acceptance probability for $\beta^{\pi c}{}_i$ is given by

$$p(accept) = \min\left(\frac{p_{Bernoulli}(B_{it} \mid \pi_{it}^{c}) * p_{normal}(\beta_i^c \mid Z_i, \gamma) * p_{poisson}(Det_{it} \mid \eta_{i,t}^{c})}{p_{Poisson}(B_{it} \mid \pi_{it}^{p}) * p_{normal}(\beta_i^p \mid Z_i, \gamma) * p_{poisson}(Det_{it} \mid \eta_{i,t}^{p})}, 1\right),$$

where π_i^c and π_i^p are obtained from Equation 4 using $\beta_i^{\pi c}$ and $\beta_i^{\pi p}$, respectively. In addition, η_{it}^c and η_{it}^p are common for $\beta^{\lambda c}{}_i$ and $\beta^{\pi c}{}_i$, as obtained from Equation 7 using β_i^c and β_i^p, respectively. We obtain an acceptance rate in the range of 20–25% with our choice of candidate distributions for β_i.

Sampling from the Conditional Distribution of ζ. We obtain draws of $\zeta = \{\zeta_0, \zeta_1, ...\zeta_{13})$ from a Metropolis-Hastings algorithm with a random walk chain. Let ζ^p denote the previous draw of ζ. The next draw ζ^c is given by $\zeta^c = \zeta^p + \breve{\zeta}$, where $\breve{\zeta}$ is a draw from a multivariate normal proposal density N (0, mΦ). We set Φ equal to the asymptotic variance–covariance matrix of ζ parameters, estimated using a two-stage model with maximum likelihood estimation. That is, we fit a Poisson regression model using Det_{it} as the dependent variable and β_i from a model specification without Equations 7 and 8. Then, m is a scalar chosen to achieve a reasonable acceptance rate. The acceptance probability is given by

$$p(accept) = \min\left(\frac{p_{poisson}(Det_{it} \mid \eta_{it}^{c}) * \exp\left[-(\zeta^c - \zeta_0)' V_\zeta^{-1} (\zeta^c - \zeta_0)\right]}{p_{poisson}(Det_{it} \mid \eta_{it}^{p}) * \exp\left[-(\zeta^p - \zeta_0)' V_\zeta^{-1} (\zeta^p - \zeta_0)\right]}, 1\right).$$

The choice of parameters for this proposal density leads to an acceptance rate of approximately 40%.

*Sampling from the Conditional Distribution of the Variance Parameters. W*e sample V_β from an inverted Wishart distribution with the degrees of freedom and scale parameter given by $(p^\lambda + p^\pi + 3) + N$ and diag(10) + $\sum_{i=1}^{N} (\beta_i - \gamma Z_i)(\beta_i - \gamma Z_i)'$, where N = number of customers.

Sampling from the Conditional Distribution of γ. We sample γ from $N(\hat{\gamma}, \hat{V}_\gamma)$, where

$\hat{\gamma} = \hat{V}_\gamma \left(V_\gamma^{-1} \gamma_0 + \sum_{i=1}^{N} Z_i' V_\beta^{-1} \beta_i \right)$, and $\hat{V}_\gamma = \left(V_\gamma^{-1} + \sum_{i=1}^{N} Z_i' V_\beta^{-1} Z_i \right)$.

APPENDIX 2

COMPUTATION OF WEIGHTED CUSTOMER ATTITUDES

The weighted attitude toward the own drug was obtained from

$$wao_i = \left(\frac{n_{i,ao}}{T}\right) * \frac{\sum_{l=1}^{n_{i,ao}} \left(ao_{i,l} / T - t_l \right)}{n_{i,ao}}, \qquad \text{(WB1)}$$

where

$n_{i,ao}$ = number of responses for own drug from customer i,

T = number of time periods in the data,

t_l = time period of customers' l^{th} response for own drug, and

$ao_{i,l}$ = customer i's attitude towards own drug in their l^{th} response.

We first weighted each response for customer attitudes toward the own drug by the time period t_l or the recency of the response. Higher importance was granted to attitude responses in later time periods, because the l^{th} attitude response is divided by a smaller number for responses closer to the end of the data time frame (T). The average of this recency-weighted attitude was multiplied by the probability of observing a response from the customer ($n_{i,ao}$/T), to account for the frequency of responses. Because customers who respond often have a higher probability of response (i.e., larger values of $n_{i,ao}$), our measure provides more weight to attitudes of customers who respond more often. Weighted attitudes toward the competing drug (wac_i) are measured similarly to wao_i and weighted relative attitude ($wrela_i$) is obtained as the difference between wao_i and wac_i.

Weighted attitudes toward the own drug for customers who never responded to the survey are predicted:

$$\hat{w}ao_i = \left(\frac{\hat{n}_{ao}}{T}\right) * \frac{\left(\hat{a}o_i / T - \hat{t}_{ao} \right)}{\hat{n}_{ao}}, \qquad \text{(WB2)}$$

where

$\hat{a}o_i$ = predicted attitudes toward the own drug for customer i, from Equations 9 and 10;

$\hat{n}_{ao}$ = average number responses among survey responders in the holdout sample;

$\hat{t}_{ao} \quad = \sum_{t=1}^{T} \frac{nresp_{t,ao}}{n_h} * t$, or the expected time period of response among survey responders in the holdout sample;

$nresp_{t,ao}$= number of responses obtained for attitudes towards own drug in time t; and

n_h = number of customers in the holdout sample.

Predicted attitudes toward the own drug for nonresponders came from the ordered probit models in Equations 9 and 10. We then weighted the predictions for expected recency ($\hat{t}$) and frequency of expected responses ($\hat{n}_{ao}$). Expectations of response recency and frequency came obtained from the empirical distribution of responses in the holdout sample. Specifically, expected frequency was the average number of responses about attitudes toward the own drug among survey responders in the holdout sample. The expected recency referred to the weighted sum of the time periods available for survey response. The weights in each time period were the ratio of the number of survey responses in each time period over the number of survey responders in the holdout sample.