An Information Stock Model of Customer Behavior in Multichannel Customer Support Services

Kinshuk Jerath†
Assistant Professor of Marketing
Tepper School of Business
Carnegie Mellon University
kinshuk@cmu.edu

Anuj Kumar
Assistant Professor of Information Systems Management
Warrington College of Business Administration
University of Florida
akumar1@ufl.edu

Serguei Netessine
The Timken Chaired Professor of Global Technology and Innovation
Technology and Operations Management Department
INSEAD
Serguei.Netessine@insead.edu

August 2012

† The authors are listed in alphabetical order and have contributed equally. They would like to thank Marketing Science Institute and the Wharton Customer Analytics Initiative for generous support for this research. They would also like to thank seminar participants at INSEAD, Indian School of Business, The Crossplatform and Multichannel Consumer Behavior conference organized by the Wharton Customer Analytics Initiative, and the Challenges and Opportunities in Managing IT Enabled Multichannel Operations conference at the University of Florida.
**An Information Stock Model of Customer Behavior in Multichannel Customer Support Services**

**Abstract**

Firms offer customer support via multiple channels, such as telephone, web portal, web chat and interactive voice-response units, but the efficacy of interactions at these channels is poorly understood. In this paper, we develop a novel information stock-based model to understand customers’ usage behavior for support services in a multichannel scenario. Our setting is that of a US-based health insurance firm. In case of a query regarding health insurance coverage or claims, customers can either call the firm’s call center to get the desired information from a call center representative, or visit the web portal and get the desired information themselves. We assume that each customer has a latent “information stock” which is a function of customers’ “information needs” (which arise when customers file health insurance claims) and “information gains” (which customers obtain when they contact the firm’s support channels to resolve their queries), and other factors such as seasonal effects (for instance, queries that arise at the time of annual contract renewal). We model a customer’s observed behavior, in terms of her query frequency and channel choice for queries, as a stochastic function of her latent information stock; this allows us to estimate the relative efficacy of different support channels.

We estimate our model on individual-customer-level data from the firm, and we find that the average information gain for a customer from a telephone call is an order of magnitude greater than that from visiting the web portal. We further find that customers prefer the telephone channel for health event-related information needs but, interestingly, prefer the web portal for seasonal information needs which are typically more structured. Additionally, information needs also vary with the nature of the health event—for instance, claims associated with repeated health events (due to, say, a chronic disease) generate minimal information needs compared to regular claims. We also find that customers are very heterogeneous in terms of their propensity to use the web channel, and can be broadly classified into “web avoiders” and “web seekers.” Besides providing the above insights, our model can be used to aid in call center management and staffing decisions as it provides very good predictions for future query volumes on different channels, and it can even help to accurately identify customers who are expected to have high telephone call volumes in the near future.

**Keywords:** multichannel customer behavior, customer service, call center, empirical OM, probability modeling.
1 Introduction

Investment into after-sales customer support is crucial for customer satisfaction and, therefore, customer retention and loyalty. The most visible example of this is a modern-day call center which, on average, accounts for approximately 70% of business-to-consumer interactions at a firm (Mandelbaum 2006). Historically, customer service representatives at call centers responded to customers’ queries over telephone. With the advent of the Internet and the World Wide Web, present day call centers offer a number of advanced technology-enabled channels to efficiently respond to customers’ queries. These support channels fall into two distinct categories: assisted channels where the firm’s representatives assist customers via telephone, email, short message service (SMS) and web chat, and self-service channels where customers can find desired information via web-based self-service portals and interactive voice response (IVR) units.

Firms have strong incentives to guide customers towards using self-service channels, as these channels cost the firm an order of magnitude lesser than assisted channels. However, customer’s channel choice will depend on the perceived values of the assisted and self-service channels, and it is not clear what these perceived values are, or how to estimate them. Furthermore, it is also important to understand the determinants of a customer’s channel choice to predict future load on different channels to optimally allocate resources at call centers. As direct labor costs account for roughly 70% of total call center costs, firms deploy a variety of statistical models to predict call traffic and thus optimally allocate manpower to contain call center costs (see Koole et al. (2003) for a review on these models). However, most of these models assume query arrival as an exogenous process and then model the service time to estimate the optimum resources at call centers, without explicitly modeling customer channel choice and interactions among different channels.

In this paper, we develop a novel information-based framework to assess the customer-perceived values of different support channels from transactional data for the support channels. We implement our model on data from the call center of a large US-based health insurance firm (name withheld due to a non-disclosure agreement) which offers web- and telephone-based support to its customers. The key idea behind the model we propose is that customers use support channels when they want to resolve queries that arise while using the product (in this case, their health insurance plan). Therefore, a customer’s observed channel usage behavior is driven by her latent information need at that time.¹

¹ Note that we use the term “information need” to refer to anything that induces the customer to contact the firm. For instance, in one case, a customer might want to understand the process to file a claim for a certain type of medical service that she has availed for the first time; this would classify as a query to gather information regarding a process she is unfamiliar with. In another case, a customer might observe that the reimbursement she received after she filed
Based on this idea, we develop an information stock-based model of customer behavior. For a customer, we assume that product usage leads to an “information need” and support channel usage leads to an “information gain” which (possibly only partially) satisfies this need. We model the latent “transactional information stock” for a customer at a given time as the sum of her information needs (which, in our context, arise when customers file health insurance claims) and information gains (which customers obtain when they contact the firm’s support channels to resolve their queries) up to that time. We assume that a customer’s observed channel usage behavior is a realization of a two-stage stochastic process—a Poisson query arrival process followed by a Bernoulli channel choice process—with the rates of the stochastic processes in both stages dependent on the information stock of the customer at that time. Besides the “transactional information stock” described above, we also account for a “seasonal information stock” to allow for queries that occur due to a time event, such as renewing the insurance contract or changes in contract terms. We further enrich our model by allowing for the information needs specific to claims associated with health events; for instance, information need may be higher if the customer has out-of-pocket expenses associated with a claim. We estimate this model on individual-customer-level data on claims and channel usage obtained from the firm. We also estimate a benchmark model without the information stock component, and find that the information stock-based model fits the data significantly better.

The parameter estimates from the information stock-based model suggest that, in our setting, the telephone channel provides, on average, an order of magnitude greater information to the customer than the web channel does. We also find that health event-related information needs vary with the nature of health events—health events where a customer has to make an out-of-pocket payment induce significantly higher information needs than health events with no payment liability from the customer, and repeated health events (say, events associated with multiple claims for a chronic disease) lead to a small fraction of the information need of the original health event. Our model estimates also provide insights regarding the determinants of channel choice. We find that customers prefer the telephone channel if health event-related information needs are higher, but prefer the web portal for seasonal information needs. Last but not the least, our model estimates the heterogeneity in customers’ baseline propensities for query frequency and web choice. Across customers, we find a high degree of heterogeneity in query propensities, and find a bipolar distribution of web choice probability indicating the existence of two distinct customer segments: “web avoiders” and “web seekers.”

---

*a claim differed from her expectation, and she might call the firm to resolve this issue; this would classify as a query for clarification. In both situations, the customer is calling the firm because, at some level, she needs information about some details of her health insurance plan.*
We also compare the in-sample and out-of-sample predictive errors of the information stock model and the benchmark (i.e., non-information stock) model. We find that information stock model provides very good predictions for total queries, telephone queries, and web queries in both in-sample and out-of-sample predictions, which are also considerably better than the predictions from the benchmark model. This finding further validates the information stock-based modeling of consumer behavior in a support services setting, and also demonstrates that our model can accurately predict the future load on different support channels at call centers. Our information stock model is also able to identify, with high precision, customers who are expected to have high probabilities of making telephone calls to the firm in the near future. This information can be extremely valuable for the firm. For instance, the firm can itself make preemptive calls to such identified customers in lean traffic periods to resolve their queries; this can help the firm to reduce some of its peak time call volume and hence reduce the customer service representative costs at call centers.

Our research makes several contributions to the literature. First and foremost, to the best of our knowledge, ours is the first attempt to formalize customer query arrival and channel choice processes by explicitly modeling a customer’s latent information stock. This model provides highly accurate predictions of future load on call centers. Second, our proposed framework utilizes information as a common denominator to understand determinants of customers’ channel usage, with different sources contributing to this latent construct. This approach allows call center managers to make inferences about the relative values of different factors that influence customers’ behavior to seek support from the firm. For the first time, we are able to utilize this approach to provide quantitative evaluation of the customers’ information demand and firms’ information supply through different channels. Third, we provide a practical framework which allows a company, using only limited transactional data on customer product and channel usage (usually captured in today’s business environment), to improve quality of service through better estimation of the query arrival process and, potentially, to even preempt query arrival.

The rest of the paper is organized as follows. In the next section, we discuss the related literature. In Section 3, we describe our research setting and our data, and conduct a preliminary analysis which supports our modeling approach. In Section 4, we develop our models, and in Section 5 we present our results. In Section 6, we conclude with managerial implications of our research and outline future research directions.

2 Related Literature

Our work is primarily related to three streams of literature: customer behavior in a service support environment, policies for call center optimization, and multichannel customer behavior.
It is of great importance for firms to understand customer behavior in support services (Sousa and Voss 2006), which has motivated several papers on this topic. Bobbitt and Dabholkar (2001) and Meuter et al. (2005) explored the determinants of adoption and customer satisfaction for self-service technology (SST) channels using questionnaires and survey tools to elicit customers’ preferences regarding SSTs. However, they did not consider how adoption of SSTs affects demand for other available alternative channels. Xue et al. (2007) show that the adoption and usage of various service channels (tellers, ATM and online banking) offered by a large retail bank depends on demographic characteristics of customers. Campbell et al. (2010) conduct a field study on the impact of online banking channel adoption on local branches, IVR, ATMs and call centers. They show that the users who adopted the online banking channel reduced their dependence on the IVR and the ATM (substitution) but increased their consumption of the firm’s call center and local branches (augmentation). Kumar and Telang (2011) conducted controlled experiments to show that the web portal is useful for providing structured information to customers, but it is not effective for resolving unstructured queries; in the latter case, they find an aggravation effect, i.e., customers who use the web portal for unstructured queries subsequently call more by telephone to the call center. In the present work, we address the question of customer channel usage by developing a probability model to estimate the relative efficacy of multiple support channels. Besides providing insights on customers’ multichannel behavior, our model also generates predictions for future activity of customers.

Call center optimization is a highly researched topic in Operations Management. However, most of this work has been done primarily using analytical queuing models (Kleinrock 1975); a comprehensive review of this research is available in Gans et al. (2003). There is a significant body of work on data-driven statistical models for predicting call center traffic to aid with staffing and workforce management decisions (Avramidis et al. 2004, Bassamboo and Zeevi 2009, Brown et al. 2005, Cezik and L’Ecuyer 2008, Mehrotra and Fama 2003, Mehrotra et al. 2010, Shen and Huang 2008, Soyer and Tarimcilar 2008, Taylor 2008, 2012, Weinberg et al. 2007). However, this literature typically models only telephone call arrivals and often assumes exogenous arrival rates for queries. In the present work, we model both query arrival and channel choice (with telephone as only one of the various channels that consumers can use to make queries), and make both of these processes endogenous to the model by making a customer’s observed behavior a stochastic function of the customer’s latent information stock, which is dependent on her history of health events and queries with the firm. Our approach is related to the recent movement in operations management towards more accurately modeling consumer behavior in operational models (Netessine and Tang 2009). Additionally, there are related empirical studies which estimate parameters
of the queuing systems in retail, healthcare and fast food settings (Olivares et al. 2012a, Olivares et al. 2012b, Pierson et al. 2011) all of which, again, only have a single stream of customers.

Recent advances in technology have enabled firms and customers to communicate via multiple channels; therefore, multichannel customer management has become one of the key challenges faced by practitioners (Neslin et al. 2006). Management of multichannel support services has become one of the hottest issues as evidenced by industry surveys, trade publications and discussions with top management; it suffices to visit such websites of solutions providers such as Multichannel Merchant and eGain. Marketing scholars have also studied issues in multichannel customer management, focusing primarily on interactions among different sales channels such as online sales, physical store sales, catalog sales, etc., (Ansari et al. 2008, Danaher et al. 2003, Deleersnyder et al. 2002, Geykens et al. 2002, Inman et al. 2004, Knox 2007, Shankar et al. 2003). In contrast, there is very little empirical work on multichannel customer support services. Sun and Li (2010) study interactions between on-shore and off-shore call centers. We do not know of any empirical work that informs how usage of self-service channels affects the usage of assisted channels in a multichannel customer support scenario. This research is an attempt to start filling this gap.

3 Research Setting, Data Description and Preliminary Analysis

We study customer behavior at a multichannel call center of a major US health insurance firm. The firm has a customer base of over three million. Customers purchase annual health insurance plans from the firm and thereafter utilize the plans to get their medical expenditures reimbursed. During the health plan usage, customers often have queries regarding their plan coverage, status of claims, etc., for which they contact the firm. During the study period, the firm offered support to its customers via telephone (assisted channel) and web portal (self-service channel). For web portal usage, customers have to first register at the web portal. Thereafter, they can visit the web portal at their convenience and obtain information regarding their plan benefits, their claim status, details of participating health providers, general information on diseases, etc. Interacting with a customer via the web portal costs the firm an order of magnitude lesser as compare to interacting via the telephone channel. This is in line with industry estimates which suggest that, per contact occasion, it costs a typical firm $0.24 on average to interact with a customer via the web portal, whereas it costs $5.50 on average to interact via telephone (Kingstone 2006). Therefore, the firm in question, like many other similar firms, is interested in: (1) understanding the determinants of customer channel choice, and (2) predicting future load on different channels for more efficient resource allocation.
We collected data for a random sample of 2462 customers from the web-registered customer population of the firm. Note that a customer’s web visit is recorded only if she logs into the firm’s website, for which she has to be registered at the web portal. Therefore, by choosing only web-registered customers, we ensure that all customers in our data have the ability to use the web channel if they so desire, and these are the customers of interest for this study.

For the 2462 customers, we constructed an individual-level dataset covering the 30-month time period from July 2005 to December 2007 by extracting relevant information from several disparate databases of the firm. Using the claims processing database of the firm, we collected data on the date of claim filing, the customer out-of-pocket expenses and the provider charges for each claim. For a customer, if a claim has the same provider charges as one of her previous claims, we term the subsequent claim as being from a repeated health event. We extracted telephone usage information, specifically, the date of a telephone call, from the Automatic Call Distributor (ACD) of the call center. Finally, we extracted web portal usage information from the web informatics database of the firm. Brief summary statistics on claims and queries are reported in Table 1.

The key idea behind our modeling approach is that filing claims leads to information needs for customers, to resolve which they contact support services with queries. This directly implies that, in our data, claims and queries should be strongly correlated. Before we develop the full model, we test for this relationship in the data through a simple analysis. Using the following fixed-effects model, we regress the total monthly number of queries for customer $i$ on the monthly number of claims she files, controlling for the other relevant variables that may affect the number of queries:

$$Q_{it} = \beta_i + \beta_t + \beta_1 \text{CLM}_{it} + \beta_2 \text{CHRG}_{it} + \beta_3 \text{RPT}_{it} + \epsilon_{it},$$
where \( i \in \{1,2,\ldots,2462\} \) denotes the customer, \( t \in \{0,1,2,\ldots,29\} \) denotes the months from July 2005 to December 2007, \( Q_{it}, CLM_{it}, CHRG_{it} \) and \( RPT_{it} \) denote, respectively, the total number of queries, total number of claims, total claim charges (in thousands of dollars) and total number of claims attributed to repeated health events in month \( t \) for customer \( i \). \( \beta_i \) is a customer-level fixed effect, and \( \beta_t \) is a month-level fixed-effect.\(^2\) The estimates of the coefficients for this regression are reported in Table 2.

\begin{table}[h]
\centering
\begin{tabular}{|c|c|c|}
\hline
& Coefficient Estimate & (Std. Err.) \\
\hline \( \beta_1 \) & 0.025*** & (0.004) \\
\hline \( \beta_2 \) & 0.003** & (0.001) \\
\hline \( \beta_3 \) & -0.014*** & (0.005) \\
\hline R-square & 0.40 & \\
\hline
\end{tabular}
\caption{Estimates for Fixed-Effects Regression}
\end{table}

\( \beta_1 \) is the coefficient of interest that specifies the net effect of the number of claims by the customer on the number of queries by the customer. We find that the value of \( \beta_1 \) is positive and highly significant. This clearly indicates that, after controlling for other factors that affect queries, and even though we expect certain amounts of noise, randomness and heterogeneity in how consumers respond to the information need generated by claims, the total number of monthly queries is strongly positively correlated with the total number of monthly claims for customers. This result gives confidence in our modeling approach. Moreover, a positive and significant value of \( \beta_2 \) indicates that queries increase for claims with higher charges, and a negative and significant value of \( \beta_3 \) indicates that queries decrease with more repeated health events. Both these effects are in the expected directions. After obtaining these preliminary reassuring results, we now proceed to develop our model.

4 Model Development

In this section, we first develop a benchmark model which treats the telephone and web channels as independent, and then we develop the information stock model. The models we build fall in the class of probability models of customer behavior (e.g., Fader and Hardie 2009). In this modeling paradigm, the

\(^2\) Repeated health events refer to multiple health events that arise due to the same health issue for the customers, e.g., multiple chemotherapy sessions. We can expect that a customer’s information needs are small for claims associated with repeated health events.

\(^3\) Month-level fixed-effects control for variations in queries with time. For instance, customers are more likely to call in certain months of the year, such as during the insurance contract renewal month or during allergy seasons.
observed behavior of a customer is modeled as a stochastic function of latent propensities of the customer.

4.1 Model 1: Benchmark Model

Invoking queuing theory-based work (Kleinrock 1975) as well as research on call center management (Gans et al. 2003), we assume that query arrival at the telephone and web channels is governed by a Poisson process. In addition, we assume that the query arrival processes at the two channels are independent. Let \( t_{ij}^T \) be the time of arrival of the \( j \)th telephone call for customer \( i \), where \( j = 1, 2, 3, \ldots, x_{iT} \) represents the sequence number of telephone calls. If \( \lambda_{i0T} \) is the baseline mean call arrival rate for customer \( i \), then the likelihood of the observed \( x_{iT} \) call arrivals for the customer is:

\[
L_{iT} = \prod_{j=1}^{x_{iT}} \lambda_{i0T} e^{-\lambda_{i0T}(t_{ij}^T - t_{i(j-1)}^T)}.
\]

Likewise, the likelihood of \( x_{iW} \) web query arrivals for customer \( i \), where the time of arrival of the \( j \)th web query is \( t_{ij}^W \), and \( \lambda_{i0W} \) is the baseline web query arrival rate for customer \( i \), is:

\[
L_{iW} = \prod_{j=1}^{x_{iW}} \lambda_{i0W} e^{-\lambda_{i0W}(t_{ij}^W - t_{i(j-1)}^W)}.
\]

For \( i = 1, 2, \ldots, N \) customers, the likelihood of the observed web and telephone queries for \( N \) customers is:

\[
L = \prod_{i=1}^{N} \left\{ \prod_{k=1}^{x_{iT}} \lambda_{i0T} e^{-\lambda_{i0T}(t_k^T - t_{k-1}^T)} \times \prod_{k=1}^{x_{iW}} \lambda_{i0W} e^{-\lambda_{i0W}(t_k^W - t_{k-1}^W)} \right\} \tag{1}
\]

In this model, \( \lambda_{i0T} \) and \( \lambda_{i0W} \) are the latent propensities governing customer \( i \)'s observed behavior. Furthermore, different customers may have different latent propensities for their query processes. Some customers may inherently have the tendency to ask more queries than others from the firm. Moreover, different customers may have different preferences for the channel to use—for instance, some customers may be very web savvy and therefore prefer the web channel over telephone, while older customers may prefer the telephone channel. We do not observe customer-level characteristics in our data and cannot control for observed heterogeneity. However, we allow for unobserved heterogeneity in customers’ behaviors by allowing the baseline telephone and web query arrival rates \( \lambda_{i0T} \) and \( \lambda_{i0W} \) to be distributed across customers as gamma distributions. Specifically, we assume that \( \lambda_{i0T} \sim \text{gamma}(\gamma_T, \theta_T) \) and \( \lambda_{i0W} \sim \text{gamma}(\gamma_W, \theta_W) \), i.e.,

\[
f(\lambda_{i0T}|\gamma_T, \theta_T) = \frac{\theta_T^{\gamma_T} \lambda_{i0T}^{\gamma_T-1} e^{-\theta_T \lambda_{i0T}}}{\Gamma(\gamma_T)} \quad \text{and} \quad f(\lambda_{i0W}|\gamma_W, \theta_W) = \frac{\theta_W^{\gamma_W} \lambda_{i0W}^{\gamma_W-1} e^{-\theta_W \lambda_{i0W}}}{\Gamma(\gamma_W)}.
\]
It is also possible to model customer activity in the following two-step manner: customers first generate queries about firm’s products/services while using it, and then choose which channel to use to resolve the query. We can model observed customer channel usage behavior by modeling query arrival for a customer as a Poisson process with the mean query arrival rate \( \lambda \) and, in the second step, modeling the customer’s choice between the web and telephone channels as a Bernoulli choice process with web choice probability \( p \). It is easy to demonstrate that, at the individual level, this model is isomorphic to the benchmark model above with a telephone query arrival rate of \( \lambda (1-p) \) and a web query arrival rate of \( \lambda p \). Therefore, we do not estimate this model separately.

### 4.2 Model 2: Information Stock Model

In the benchmark model, the query arrival processes for both channels are assumed to be exogenous. However, we wish to understand the underlying mechanism that drives the query arrival and channel choice processes for customers. An effective approach to do this, used widely in the marketing and economics literatures, is by modeling a latent construct that drives observed behavior. For instance, McFadden (1973) introduced the idea that the latent construct, “utility,” drives the stochastic choice process leading to observed consumer choice. Moe and Fader (2004) examine consumers’ dynamic purchase behavior at an e-commerce website with a latent visit effect that evolves over visits. Netzer et al. (2008) examine alumni gift-giving behavior by using a nonhomogeneous hidden Markov model in which donors transition from one latent relationship state with their university to another.

In a similar vein, our approach is to construct a plausible model of the underlying mechanism that leads to customers’ observed multichannel behavior. We assume that, at any point in time, each customer has an “information stock” which determines her query frequency and channel choice behaviors. The information stock of a customer, in turn, is determined by the information needs that arise about the insurance contract as she uses the insurance plan (for instance, queries about their medical coverage, claim processing procedures, and operational details of the contract), and the information gain that she obtains upon contacting the firm through the telephone or the web channels.

We categorize a customer’s information stock into two broad categories, as appropriate for the setting we model. The first category is transactional information stock, which is determined by the health events faced by a customer. We assume that each insurance claim filed by the customer (or by a doctor’s office on the customer’s behalf) after a health event leads to an information need \( C \) for the customer. In other words, with each additional claim, the information need of the customer increases by \( C \). In order to fulfill her information needs, the customer can approach the firm through its support channels and receive some information gain from contacting the support channels. We assume that each web portal visit
provides the customer with an information gain $W$, and each telephone call provides the customer with an information gain $T$. These information needs and gains determine her transactional information stock. Note that transactional information stock varies across customers based on the number of claims they file and the queries they make.

The claims filed and queries made by customers over a period can be organized into a set of sequences, where each sequence contains several claims followed by a query. Consider a customer denoted by $i$. Assume that this customer makes $x_i$ queries with $t_{ij}$ as the time of the $j^{th}$ query, $j=0, 1, 2, 3, \ldots, x_i$. Between two queries, several claims are filed by the customer. Let $n_{ij}$ be the number of claims that are filed between the $j^{th}$ and $(j+1)^{th}$ query by customer $i$. We denote the net transactional information stock after the $j^{th}$ query for customer $i$ by $I_{ij}$, given by:

$$I_{ij} = \sum_{z=0}^{j} [n_{iz}C - (y_{iz}W + (1 - y_{iz})T)],$$

where $y_{iz} = 1$ if the $z^{th}$ query for customer $i$ is a web portal visit and 0 otherwise.

Furthermore, let $t_{ijk}$ denote the time of arrival of the $k^{th}$ claim after the $j^{th}$ query with $k=0, 1, 2, 3, \ldots, n_{ij}$. We denote the net transactional information stock for customer $i$ at time $t_{ijk}$ (i.e., after the $k^{th}$ claim after the $j^{th}$ query) by $I_{ijk}$, given by:

$$I_{ijk} = I_{ij} + kC.$$

The second category of information stock is seasonal information stock, which is determined by needs and gains that arise due to insurance plan-related events in time. For instance, around the time of insurance contract renewal, customers make more queries regarding their ID card, renewal of their web portal login and password, insurance forms to be used in the upcoming year, etc. Similarly, the firm sends seasonal information bulletins to consumers, in allergy seasons, for instance, leading to information gains. These seasonal information needs are variable with time but applicable to all customers at a given point in time. We model seasonal information needs at the monthly level. We use 24 months of data to estimate our model. We use parameters $I_{m}$ to denote the seasonal information stock, equally applicable to all customers, in month $m$, where $m \in \{0, 1, 2, \ldots, 23\}$.

The total stock of information need for customer $i$ after the $k^{th}$ claim after the $j^{th}$ query, and when the month is $m$, is given by $I_{ijk} + I_{m}$, and is the sum of the transactional and seasonal information stocks, with the total stock being zero at time zero. Note that $C$, $W$ and $T$ are assumed to be the same for all claims, web visits, and telephone calls, respectively. The dynamics in the total information stock for a customer are shown in Figure 1.

Observed customer query arrival and channel choice are driven by the total information stock that the customer has at a given time. We model the query arrival process as a Poisson process. The rate for this process, for customer $i$ after the $k^{th}$ claim after the $j^{th}$ query, and when the month is $m$, is given by:
\[ \lambda_{ijkm} = \lambda_{i0} \exp(l_{ijk} + l_m), \tag{2} \]

Figure 1: Dynamics in a Customer’s Information Stock over Time (\(m1\) and \(m2\) denote Month 1 and Month 2, respectively)

where \(\lambda_{i0}\) is the baseline mean query arrival rate for customer \(i\). Therefore, the mean query arrival rate, \(\lambda_{ijkm}\), will change for a customer with arrival of claims or queries or with a change in month, i.e., query arrival for a customer is modeled as a nonhomogeneous Poisson process with the mean rate of arrival changing with the customer’s information stock. Note that a higher information stock denotes a higher information need, which corresponds to a larger query arrival rate.

We assume that, after the customer has decided to make a query, she makes a Bernoulli choice between using the web and making a telephone call to resolve her query, with a web choice probability \(p\). Since telephone calls are answered by trained representatives of the firm, it is likely that when the information need is high, customers prefer making a telephone call. Likewise, it may be possible that customers prefer the web portal for structured information needs such as seeking insurance contract-related information, such as applying for a new ID card. We allow for these possibilities by modeling the web choice probability, \(p_{ijm}\), as a function of the two types of information stocks. We define the web choice probability for the \(j^{th}\) query for customer \(i\), where the \(j^{th}\) query for the customer arrives in month \(m\), as:
\[ p_{ijm} = p_{i0} \frac{\exp(\pi_T I_{ij} + \pi_S I_{jm})}{[1 - p_{i0} + p_{i0} \exp(\pi_T I_{ij} + \pi_S I_{jm})]}, \quad (3) \]

where, \( p_{i0} \) indicates customer \( i \)'s baseline web choice probability independent of information need. The parameters \( \pi_T \) and \( \pi_S \) are included to make the model flexible by allowing the impacts of the transactional and seasonal information stocks, respectively, to be different on the channel choice probability than on the query arrival rate. The values that these parameters take inform us on the sensitivity of the channel choice probability to the two kinds of information stocks. Note that the Bernoulli web choice probability \( p_{ijm} \) for customer changes with the arrival of claims and queries as well as with the change of month. (Note that we need channel choice probability only at the time of a query, and not after every claim, which is why we do not need to define \( p_{ijklm} \), where \( k \) indexes claims between queries.)

We are now ready to develop the likelihood function for the observed data for customer \( i \). For notational simplicity, we suppress the subscript \( m \) for month in the following derivation, i.e., we ignore the impact of the seasonal information stock and focus on the impact of the transactional information stock. Incorporating the seasonal information stock via month is straightforward since we only have to modify the query arrival rate and channel choice probability based on the month at a specific time. The customer receives \( n_{i0} \) claims \((1, 2, \ldots, n_{i0})\) at time \( t_{i01}, t_{i02}, \ldots, t_{i0n_{i0}} \) and then the first query arrives at time \( t_{i1} \). The processes up to the first query of the customer are:

No query up to \( t_{i01} \) rate \( \lambda_{i00}; \) no query between \( t_{i01} \) and \( t_{i02} \) at rate \( \lambda_{i01}; \ldots \); no query between \( t_{i0(n_{i0}-1)} \) and \( t_{i0n_{i0}} \) at rate \( \lambda_{i0(n_{i0}-1)} \); query arrival at \( t_{i1} \) with rate \( \lambda_{i0 n_{i0}} \) between \( t_{i0 n_{i0}} \) and \( t_{i1} \); channel choice by customer for first query with web choice probability \( p_{i1} \).

Therefore, the likelihood function for customer \( i \) up to the first query is given by:

\[
L_{i1} = e^{\sum_{u=0}^{n_{i1}-1} \lambda_{i0 u} (t_{i0(u+1)} - t_{i0 u})} \times \lambda_{i0 n_{i1}} e^{\lambda_{i0 n_{i1}} (t_{i1} - t_{i0 n_{i1}})} \times p_{i1} y_{i1} (1 - p_{i1})^{(1 - y_{i1})},
\]

where \( y_{i1} = 1 \) if the first query for customer \( i \) is a web portal visit and \( 0 \) otherwise.

The customer receives a total of \( x_i \) queries with \( n_z \) claims between \( z^{th} \) and \( (z+1)^{th} \) query, where \( z=0, 1, 2, \ldots, (x_i-1) \). The customer \( i \) receives \( g_i \) claims after the \( x_i^{th} \) query till the end of our period of observation \( (t_{end}) \), which is the same for all customers. The total likelihood function for customer \( i \) is given by:

\[
L_i = \prod_{z=0}^{x_i-1} \left[ e^{\sum_{u=0}^{n_{iz}-1} \lambda_{iz u} (t_{iz(u+1)} - t_{iz u})} \times \lambda_{iz n_{iz}} e^{\lambda_{iz n_{iz}} (t_{i(z+1)} - t_{iz n_{iz}})} \times p_{i(z+1)}^{y_i(z+1)} (1 - p_{i(z+1)})^{(1 - y_i(z+1))} \right] \times e^{\sum_{u=0}^{g_i} \lambda_{ix u} (t_{ix(u+1)} - t_{ix u})} \quad (4)
\]
where $t_{i(k+1)} = t_{\text{end}}$.\(^4\)

If there are $i=1, 2, \ldots, N$ customers, the likelihood of observed web and telephone queries for $N$ customers is given by:

$$L = \prod_{i=1}^{N} L_i$$

(5)

As the subscript $m$ for month was suppressed in (4), the Poisson mean query arrival rate $\lambda_{ijk}$ and the Bernoulli web choice probability $p_{ij}$ utilized in (4) and (5) are actually $\lambda_{ijkm}$ and $p_{ijm}$ computed with appropriate transactional and seasonal information stock components as per Equations (2) and (3), respectively.

So far we have assumed that all claims for a customer give her the same information need $C$. However, claims with different characteristics may lead to different information needs. For instance, customers are more likely to make queries for claims where they have to pay out-of-pocket charges or for claims of higher value. We also find that customers often face repeated health events of similar nature, e.g., multiple chemotherapy or dialysis sessions. In such cases, the customers may have higher information needs in the first few claims but once they understand their insurance plan coverage, they may have little to ask in further repeated claims. Therefore, we allow for different information needs for claims based on whether the customer has to pay out of her pocket and whether the particular health event is a repeated health event.\(^5\) For customer $i$’s claim associated with health event $h$, we assume that:

$$C_{ih} = C_0 \exp(\alpha_{\text{LIAB}} \cdot D_{\text{LIAB},ih} + \alpha_{\text{RPT}} \cdot D_{\text{RPT},ih}),$$

(6)

where, $D_{\text{LIAB},ih}$ is a dummy variable which is equal to 1 if the claim has positive customer out-of-pocket expenses and 0 otherwise, $\alpha_{\text{LIAB}}$ is the impact of a claim with positive customer liability on the information need created by the claim, $D_{\text{RPT},ih}$ is a dummy variable which is equal to 1 if the claim pertains to a repeated health event and 0 otherwise, $\alpha_{\text{RPT}}$ is the impact of a repeated health event’s claim on the information need created by the claim, and $C_0$ is the baseline information need from a claim which is constant across all claims and across all customers.

It is natural to assume that customers may be different in their propensities to make queries and to use the web portal. Therefore, across customers, we allow for gamma-distributed heterogeneity in the

\(^4\) Note that, at the time of estimation, we use the values $C \times 10^3$, $W \times 10^3$ and $T \times 10^3$ to prevent the exponential functions in the likelihood expression from taking values that are too large. This, however, does not change the relative impact of $C$, $W$ and $T$ on the customers’ behavior.

\(^5\) We did not include the claim values here, as they are highly correlated to the customer out-of-pocket expenses.
baseline query arrival rate, \( \lambda_{i0} \sim \text{gamma}(\gamma, \theta) \), and beta-distributed heterogeneity in the baseline web-choice probability, \( p_{i0} \sim \text{beta}(a, b) \), i.e.,

\[
f(\lambda_{i0} | \gamma, \theta) = \frac{\theta \lambda_{i0}^{(\gamma-1)} e^{-\lambda_{i0} \theta}}{\Gamma(\gamma)} \quad \text{and} \quad f(p_{i0} | a, b) = \frac{p_{i0}^a (1-p_{i0})^b}{\beta(a, b)}.
\]

Finally, we recognize that the information gains from different web visits and telephone calls may be different for different customers, but given the limited data that we have, here we estimate the average information gains (\( W \) and \( T \)) from these channels for our sample of customers. Note that we allow for negative values of \( W \) and \( T \), i.e., we allow for the possibility that, on average, rather than providing information to customers, a channel aggravates the information need of customers.

Note that we did not have access to demographic information for the customers, the content of any telephone calls made by a customer, or the web pages visited by a customer during a web visit. Our data are primarily on the “transactions” of the customer with the firm; the main reason for using these data is that such data are easy to collect and have no related privacy issues, unlike customer demographic information. Since most firms have ready access to such data, we expect our model to be useful for a typical firm’s customer support center. Moreover, even with the limited data that we use, our model demonstrates excellent predictive power, as will become evident shortly.

**Identification.** We now briefly discuss how the parameters in the proposed model are identified given the variation in our data. First, the set of parameters \( I_m \), which determine seasonal information stock in different months, are identified by the common variation in query rates and web-choice rates across calendar months for all customers. The differences in baseline query and web-choice propensities across customers are identified by the differences in their overall mean query arrival rates and web-choice rates respectively after accounting for the common monthly variations. The parameters \( \gamma \) and \( \theta \), which determine the heterogeneity distribution in the baseline query propensities, are identified by the assumed parametric form of distribution on differences in overall mean query arrival rates across customers. Similarly, the parameters \( a \) and \( b \), which determine the heterogeneity distribution in the baseline web-choice propensities, are identified by application of the assumed parametric form of distribution on the differences in overall web-choice rates across customers (Note, however, that we use flexible distributions that can take various shapes.) The transactional information stock parameters \( C_0, W \) and \( T \) are identified by the variation in the query rates at different channels for different periods with respect to the average overall claim arrival and query rates for customers in that period, after controlling for the baseline query rates for the customers and common monthly variations in query rates across customers. The parameters \( a_{LIA} \) and \( a_{RPT} \), which determine the impact on information need generated by claims for
which customers have out-of-pocket expenses and claims for repeated health events, are identified by the
variation in the respective observable characteristics across different claims. The parameters $\pi_T$ and $\pi_S$, which determine how channel choice is influenced by the transactional and seasonal information stocks, respectively, are identified by the variation in web-choice probabilities with the changes in the two categories of information stocks, after controlling for overall channel choice probabilities.

5 Estimation and Results

5.1 Estimation Procedure

For the information-stock model (Model 2), we estimate two sets of parameters: (1) information stock-related parameters ($C_0$, $W$, $T$, $\alpha_{LAB}$, $\alpha_{RPT}$, $I_m$, $\pi_T$, $\pi_S$), and (2) parameters that determine heterogeneity in baselines rates across customers ($\gamma$, $\theta$, $a$, $b$). For our benchmark model (Model 1) we only estimate the heterogeneity parameters ($\gamma_T$, $\theta_T$, $\gamma_W$, $\theta_W$). For both models, we can infer individual-level parameters that vary across customers ($p_{io}$ and $\lambda_{io}$) from the estimated parameters and data on the individual’s observed behavior.

We use a hierarchical Bayes framework for model estimation (Gelman et al. 2009), using the following Markov Chain Monte Carlo (MCMC) chains (details are provided in the appendix):

- Draw ($\lambda_{io}$, $p_{io}$ | $C_0$, $W$, $T$, $I_m$, $\pi_T$, $\pi_S$, $\alpha_{LAB}$, $\alpha_{RPT}$, $\gamma$, $\theta$, $a$, $b$, data) using the Metropolis-Hastings algorithm;
- Draw ($C_0$, $W$, $T$, $I_m$, $\pi_T$, $\pi_S$, $\alpha_{LAB}$, $\alpha_{RPT}$ | $\gamma$, $\theta$, $a$, $b$, $\lambda_{io}$, $p_{io}$ data) using the Metropolis-Hastings algorithm;
- Draw ($\gamma$, $\theta$, $a$, $b$ | $C_0$, $W$, $T$, $I_m$, $\pi_T$, $\pi_S$, $\alpha_{LAB}$, $\alpha_{RPT}$, $\lambda_{io}$, $p_{io}$) using the Metropolis-Hastings algorithm.

Parameter Recovery. We have already argued that the parameters of the information stock-based model are well-identified given the data we use to estimate the model. Additionally, to check whether the proposed model can accurately recover parameter values from data, we conducted a simulation study. In this study, we simulated data from the model using sets of pre-determined parameter values to cover a variety of cases with respect to the relative values of the different support channels and the heterogeneity in latent propensities across customers. Then, using the procedure described above, we estimated the model on the simulated data to check: (i) whether the recovered parameter values match the actual parameter values used for data generation, and (ii) whether the estimated parameter values can accurately recover aggregate query volumes in the data. We find that, in all the cases that we considered, recovery of parameters as well as of aggregate statistics in the data is very good. This provides further confidence in our model. More details are available in the appendix.
5.2 Results

We calibrated Model 1 and Model 2 on the first 24 months of data (from July 2005 to June 2007), and used the last six months of data (July 2007 to December 2007) as a hold-out sample. For each model, we ran 40,000 iterations of the MCMC steps; the first 30,000 iterations were used as initial burn-in to reach convergence, which we checked visually, and the last 10,000 iterations were used to infer the posterior distributions of the parameters. We used multiple starting values for the MCMC chains and confirmed that the parameters converged to the same values. We report the estimation results in Table 3. We report the values of the 23 seasonal information stock parameters in Table A1 in the appendix.

The higher value of log marginal density for the information stock model and the large value of the log Bayes factor suggest that the information stock model fits the observed data better than the benchmark model. The heterogeneity plot of mean query arrival rate across customers in our sample is reported in Figure 2(a). This long-tailed plot indicates that there is large heterogeneity in the baseline query arrival rates—a majority of the customers have a low query rate while a few customers have the propensity to make large numbers of queries. The median number of queries is 1.84 per year.

The distribution of web choice probability across customers in our sample is reported in Figure 2(b), and indicates a polarized distribution, i.e., some customers (a relatively larger number) have quite low web-choice probability and can be classified as “web avoiders,” whereas other customers (a relatively smaller number) have quite high web-choice probability and can be classified as “web seekers.” The median web choice probability is 0.38.
<table>
<thead>
<tr>
<th></th>
<th>Model 1 (Benchmark Model)</th>
<th>Model 2 (Info. Stock Model)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>In-Sample Fit Statistics</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Log Marginal Density</td>
<td>-117131.9</td>
<td>-115203.2</td>
</tr>
<tr>
<td>Log Bayes Factor</td>
<td></td>
<td>1928.7</td>
</tr>
<tr>
<td><strong>Parameter Estimates</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\gamma_T$</td>
<td>0.860</td>
<td>0.283</td>
</tr>
<tr>
<td></td>
<td>[0.764, 0.905]</td>
<td>[0.272, 0.306]</td>
</tr>
<tr>
<td>$\theta_T$</td>
<td>175.982</td>
<td>32.126</td>
</tr>
<tr>
<td></td>
<td>[155.55, 189.91]</td>
<td>[30.099, 36.108]</td>
</tr>
<tr>
<td>$\gamma_W$</td>
<td>0.283</td>
<td>0.470</td>
</tr>
<tr>
<td></td>
<td>[0.272, 0.306]</td>
<td>[0.431, 0.517]</td>
</tr>
<tr>
<td>$\theta_W$</td>
<td>32.126</td>
<td>0.658</td>
</tr>
<tr>
<td></td>
<td>[30.099, 36.108]</td>
<td>[0.616, 0.701]</td>
</tr>
<tr>
<td>$\alpha_LAB$</td>
<td>0.583</td>
<td>2.694</td>
</tr>
<tr>
<td>$\alpha_RPT$</td>
<td>0.583</td>
<td>2.694</td>
</tr>
<tr>
<td>$\pi_T$</td>
<td>-3.930</td>
<td>-3.930</td>
</tr>
<tr>
<td></td>
<td>[-5.249, -3.014]</td>
<td>[-5.249, -3.014]</td>
</tr>
<tr>
<td>$\pi_S$</td>
<td>1.590</td>
<td>1.590</td>
</tr>
<tr>
<td></td>
<td>[1.388, 1.792]</td>
<td>[1.388, 1.792]</td>
</tr>
<tr>
<td>Table 3: Estimation Results (for each parameter, we report the posterior mean followed by the 95% credible interval)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
The distribution of web choice probability across customers in our sample is reported in Figure 2(b), and indicates a polarized distribution, i.e., some customers (a relatively larger number) have quite low web-choice probability and can be classified as “web avoiders,” whereas other customers (a relatively smaller number) have quite high web-choice probability and can be classified as “web seekers.” The median web choice probability is 0.38.

The estimated value of the baseline information need from a claim, $C_0$, is 2.694, of the information gain from a web visit, $W$, is 0.658, and of the information gain from a telephone call, $T$, is 20.768. To obtain a more practical sense of how these values compare with each other, it is helpful to analyze their impact on observable metrics. One such metric is the expected time of the next query by a customer. In terms of this metric, we find that, for a representative customer, the information gain from one telephone call is sufficient to meet the information need that arises from approximately eight claims. Similarly, the information gain from one telephone call is roughly equivalent to the information gain from 32 web portal visits.  

The above analysis leads to the interesting implication that, in our setting, a telephone call, on average, provides a large information gain to the customer in comparison to other channels—an order of magnitude higher than the information need from a claim and almost two orders of magnitude higher than the information gain from a web visit. This result also shows that, in our setting, web visits significantly less effective than telephone calls in resolving queries. This, in fact, is very much in line with the expectations of the firm’s managers whom we briefed on the results of this paper.

Next, we examine the impact of the two types of information stocks on the probability that a customer uses the web channel to resolve a query. The negative and significant estimate for $\pi_T$ (-3.930) suggests that the probability of web usage decreases with higher transactional information stock. Since transactional information stock is generated from health events, this implies that customers prefer to use the telephone channel for information needs generated by health events. In contrast, the positive and

---

6 We conduct the analysis to compare information need from a claim and information gain from a telephone call as follows. Assume that there is a representative customer who has zero information stock, i.e., only her baseline propensities are driving her query behavior, and the expected time until her next query is $D$ days. We (artificially) endow her with the information that she would need after $C$ claims, i.e., her information stock is $zC$. Suppose this reduces the expected time of her next query to $D-d_1$ days. We now (again, artificially) endow her with the information gain that she would obtain from one telephone call, $T$, so that her information stock is $zC+T$. Suppose this gain increases the expected time of her next query to $D-d_1+d_2$ days. We solve for the value of $z$ for which $D-d_1+d_2=D$, i.e., $d_1=d_2$; given our parameter estimates, this value of $z$ is approximately eight. Similarly, to compare the information gain from a web visit and a telephone call, suppose the expected time of the next query of a representative hypothetical customer starting with zero information stock and then endowed with the information gain from one telephone call, $T$, is $D-d_1$ days. We solve for the value of $zW$ such that the expected time of the next query of a representative customer starting with zero information stock and then endowed with the information gain from $zW$ web visits, $zW+W$, is also $D-d_1$ days. This value of $zW$ is approximately 32.
Table 4: Prediction Errors (MAPE)

<table>
<thead>
<tr>
<th></th>
<th>In-Sample Predictions</th>
<th>Out-of-Sample Predictions</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Total queries</td>
<td>Telephone queries</td>
</tr>
<tr>
<td>Model 1 (Benchmark Model)</td>
<td>15.06%</td>
<td>8.63%</td>
</tr>
<tr>
<td>Model 2 (Info Stock Model)</td>
<td>2.22%</td>
<td>7.01%</td>
</tr>
</tbody>
</table>

significant estimate for $\pi_S$ (1.590) suggests that the probability of web usage increases with higher seasonal information stock. As pointed out earlier, seasonal information needs are typically related to the insurance contract, such as requests for ID cards, password/login updating, etc. This category of information is easy to retrieve on the web portal and thus customers tend to use the web portal for obtaining this type of information. These results are also in line with the results of Kumar and Telang (2011), who also studied a health insurance setting and found the web to be effective in resolving structured queries, but found the telephone to be effective in resolving unstructured and complex queries.

The positive and significant estimate of $\alpha_{LIA}$ (0.583) indicates approximately 79% higher information need from claims where customers have out-of-pocket expenses as compared to claims with no customer out-of-pocket expenses. A negative and significant estimate of $\alpha_{RPT}$ (-11.944) indicates that a repeated health event generates less than 0.1% of the information need generated by the original health event.

Taking an overall view, the insights obtained from our parameter estimates lend face validity to our framework. Note that the values above are specific to this setting and depend on many factors, such as the level of training of the customer service representatives who take calls and the helpfulness and ease-of-use of the website of the company. Needless to say, these values may differ in other settings and our model will have to be re-estimated to be informative for those settings.

5.3 Model Predictions

For both of the estimated models, we predict total queries, telephone queries and web queries for each customer in our sample for the calibration period (July 2005 to June 2007) as well as the hold-out period (July 2007 to December 2007). To test these predictions from the models, we aggregate them across the full cohort of customers, and also aggregate them across time to the monthly level. We report the Mean Absolute Percentage Error values for the in-sample and out-of-sample predictions from the two models in Table 4. It is clear from this table that, as compared to the benchmark model, the information stock model
makes significantly superior in-sample and out-of-sample predictions for total queries, telephone queries, as well as web queries.

Furthermore, in Figure 3, we graphically compare the month-wise in-sample predictions of queries from the information stock model and the benchmark model with the actual values. It is apparent from Figure 3 that the information stock model tracks the actual number of total queries, telephone queries and web queries very well, and significantly better than the benchmark model. Note that the benchmark model does an excellent job of capturing the mean level of activity, but it is unable to capture temporal variations in activity.

Additionally, the information stock model can predict queries at the individual-customer level based on each customer’s calculated information stock. Using these predictions, we can use the model to better identify, as compared to the benchmark model, the customers who are likely to make a telephone call in a particular time period (say, one month). To test how well the model can do this, we compute the calling probability for each customer for each month in the out-of-sample period (July 2007 to December 2007), and then sort customers in descending order of their calling probabilities given by the information stock model. We do the same for the benchmark model as well. We then consider the top 20% highest-calling-probability customers from both models, and compute the number and percentage of correctly identified calling customers (based on the actual outcomes). We report these results in Table 5.

The results in Table 5 clearly show the impressive performance of the information stock model—it is able to correctly identify a large percentage of calling customers, which is also much higher as compared to the benchmark model. This feature of our model can be a great advantage to the firm. For instance, once the high-calling-probability customers are identified, to resolve their queries pro-actively and pre-empt some of their calls, the firm’s customer service representatives (CSRs) can make outgoing calls to them in non-peak times of the day when the CSRs are free. In other words, information about calling customers can be utilized to reduce peak-time telephone calls by using interventions of the above nature. As the CSRs are deployed at call center primarily based on the predicted peak-time call load, making calls to customers in advance may reduce the peak-time call load and thus save CSR-related costs for the company. For instance, the information stock model, on average, is able to correctly identify approximately 76% of the calling customers in a month (if the top 20% of predicted customers are considered). If 30% of the total customers make calls at the peak time, then approximately 23% of such identified customers would have made calls at the peak time. If we assume that by making outgoing calls to these customers, half of these calls are avoided, i.e., approximately 10% of total calls are avoided, it would result in significant cost savings for the firm. For instance, the firm we obtained data from has
Figure 3: Comparison of Monthly Predicted and Actual Queries
Table 5: Identification of Calling Customers

<table>
<thead>
<tr>
<th>Month</th>
<th>No. of customers making telephone calls</th>
<th>No. of customers correctly identified</th>
<th>% of customers correctly identified</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Model 1</td>
<td>Info stock model</td>
</tr>
<tr>
<td>Jul-07</td>
<td>277</td>
<td>130</td>
<td>199</td>
</tr>
<tr>
<td>Aug-07</td>
<td>286</td>
<td>121</td>
<td>209</td>
</tr>
<tr>
<td>Sep-07</td>
<td>247</td>
<td>106</td>
<td>191</td>
</tr>
<tr>
<td>Oct-07</td>
<td>277</td>
<td>103</td>
<td>222</td>
</tr>
<tr>
<td>Nov-07</td>
<td>250</td>
<td>105</td>
<td>194</td>
</tr>
<tr>
<td>Dec-07</td>
<td>206</td>
<td>87</td>
<td>156</td>
</tr>
</tbody>
</table>

approximately three million customers making on average 440,000 calls per month; in this case, making outgoing calls to the identified customers may result in a reduction of roughly 44,000 peak-time calls per month, or 2,000 peak time calls per day (assuming 22 working days per month), which can lead to very significant savings.

Call centers typically make staff allocation plans at the weekly, daily or, sometimes, even hourly levels. Even though our analysis is conducted at the monthly level, the results obtained can be used to make decisions for shorter time frames. For instance, once a planner is given a list of customers with a high probability of calling in the near future (specifically, in the next one month), he can use this information to make scheduling and staff allocation decisions for time frames shorter than one month, as the need may be. Naturally, if we had more data, we could model information stock not at the monthly but at the weekly or even daily levels, which is expected to improve the predictive power of the model. Since we are limited to a data sample, we model information stock at the monthly level; however, we conducted a weekly analysis as well and obtained insights and predictive performance qualitatively similar to what is presented here.

Finally, we conducted various robustness checks and find that there are no qualitative differences in the parameter values and insights obtained. For instance, we calibrated the model on the first 18 months of data as well as the full 30 months of data (instead of 24 months of data) and obtained similar parameter estimates. We also estimated a model with 11 variables for \( I_n \) to capture seasonal information stock (one for each calendar month across years, i.e., the same parameter for July 2005 and July 2006, for August 2005 and August 2006, and so on), and obtained similar estimates to those reported here for the other parameters. We also checked for possible learning by consumers for the web channel by allowing for higher information gains in subsequent web visits by using a multiplicative factor; we did not find this effect and, therefore, dropped this component from the final specification.
6 Conclusions

We propose a novel information stock-based framework to endogenously model the query generation and channel choice processes of customers in a multichannel customer support services setting. We postulate that each customer has a latent information stock which determines her behavior. The information stock is determined by information needs which are generated by insurance claims corresponding to customers’ health events, and information gains that customers obtain on contacting the firm’s support center by the telephone or the web channel. This information stock-based model allows us to use observed customer data—namely the sequence of claims and the query behavior—to estimate the customer-perceived information values of different support channels. We thus contribute to the Operations Management literature on call center management by providing a methodology which relates the query arrival and support channel choice processes to customer transaction history, rather than treating the query formation process as exogenous, as is typical in queuing literature.

We implement the proposed model on individual-customer-level data obtained from a large US-based health insurance firm. In this setting, we find that the average information gain from a telephone call is an order of magnitude greater than the information gain from visiting the web portal, i.e., the telephone channel, on average, is significantly more effective than the web channel in resolving customers’ queries. We find that information needs from claims vary with the nature of the health events associated with the claims—a health event whereby a customer has to make an out-of-pocket payment leads to higher information need than a health event with no payment liability from the customer, and repeated health events lead to negligible information need compared to the original health event. Regarding channel choice, we find that customers prefer the telephone channel for higher health event-related information needs while they prefer the web portal for more structured seasonal information needs. We also find that there is a large degree of heterogeneity in customers’ propensity to use the web portal while making a query—some customers are “web avoiders” while others are “web seekers.”

In addition to providing the above insights, the information stock model provides accurate predictions for aggregate query volumes for the different support channels. Furthermore, it is able to identify with high accuracy the customers who are likely to make queries in the near future. The model therefore can serve as a useful managerial tool. For instance, making advance outgoing calls to customers who have high calling probability can help to reduce peak-time calls, which can lead to substantial cost savings for the firm.

Taking a more strategic perspective, our results have significant managerial implications for new-generation multichannel customer service operations technologies. At a time when web portals are
becoming a popular choice as a way to reduce the cost of customer service, our results show that there are more subtle and complex phenomena at play. Our estimates indicate the superior informational value of the traditional telephone channel over the self-service web portal. This suggests that the assisted telephone channel is still a dominant customer support channel at least for complex services such as health insurance. Our estimates also suggest that web portals are effective for simple, unambiguous tasks (such as seasonal information needs, which are more structured and routine information needs). We expect these results to extend to other self-service support channels. Therefore, managers have to make a balanced and careful choice regarding what infrastructure they should set up for customer support services. The design of the web portal, in terms of ease of access of information, can be an important dimension in this decision.

The present model can be extended in many different directions in the future. First, we have conceptualized information as a one-dimensional construct. It may be useful to think of information as a multi-dimensional construct, e.g., one dimension can capture structured information needs and another can capture unstructured information needs. Identification of the parameters of such a model would require richer data than we have used here. Second, we do not model the claim arrival process for customers. Future work could add this component to the model, which could also help to model endogeneity between claim frequency and query behavior. Third, as appropriate for our setting, we have two channels in the model, namely telephone and web. Future work can extend the model to include more customer support channels, as the need may be; this would be a fairly straightforward extension. Fourth, we have timing data on the claims and queries of customers, and some data on claim characteristics. On the one hand, this implies that our model is useful for the typical firm since these data requirements are not heavy while the model still generates useful predictions and insights. On the other hand, these data limitations present an opportunity to further enrich the model by extending it to incorporate more data, such as the transcripts of calls made by customers and the web pages visited by customers during a web portal visit. We hope future research builds and improves on the information stock framework proposed in this paper.
References


Appendix: Details of Estimation Procedure

We recursively generate parameter draws using the MCMC chains as below:

- **Generate** $\lambda_{\ell0}$

  Posterior $\{\lambda_{\ell0} | p_{\ell0}, C_{\ell0}, W, T, \gamma, \theta, a, b, I_m, \pi_T, \pi_S, \alpha_{LAB}, \alpha_{RPT}, \text{data}\}$
  \[
  \propto L(\text{data}) \times \text{gammapdf}(\lambda_{\ell0} | \gamma, \theta) \tag{A1}
  \]
  As (A1) does not have a closed form, the Metropolis-Hastings (MH) algorithm was used to draw from the conditional distribution of $(\lambda_{\ell0})$. The $(t+1)^{\text{th}}$ draw of $(\lambda_{\ell0})$ is given as
  \[
  (\lambda_{\ell0})^{(t+1)} = \text{Truncated Normal random number} \{ (\lambda_{\ell0})^{(t)}, \text{Var} (\lambda_{\ell0}) \},
  \]
  where the normal distribution is truncated below at 0, and Var $(\lambda_{\ell0})$ is adaptively chosen to reduce the autocorrelation among the MCMC draws with an acceptance rate between 20% and 30%. The probability of accepting $(\lambda_{\ell0})^{(t+1)}$ is
  \[
  \text{Min} \left\{ \frac{(L(\text{data})^{t+1} \times \text{gammapdf}((\lambda_{\ell0})^{(t+1)} | \gamma, \theta))}{(L(\text{data})^t \times \text{gammapdf}((\lambda_{\ell0})^t | \gamma, \theta))} \times \frac{\text{Tr-Normpdf}((\lambda_{\ell0})^{(t+1)} | (\lambda_{\ell0})^t, \text{Var}(\lambda_{\ell0}))}{\text{Tr-Normpdf}((\lambda_{\ell0})^t | (\lambda_{\ell0})^t, \text{Var}(\lambda_{\ell0}))}, 1 \right\}. \]

- **Generate** $p_{\ell0}$

  Posterior $\{p_{\ell0} | \lambda_{\ell0}, C_{\ell0}, W, T, \gamma, \theta, a, b, I_m, \pi_T, \pi_S, \alpha_{LAB}, \alpha_{RPT}, \text{data}\}$
  \[
  \propto L(\text{data}) \times \text{betapdf}(p_{\ell0} | a, b) \tag{A2}
  \]
  As (A2) does not have a closed form, the MH algorithm was used to draw from the conditional distribution of $(p_{\ell0})$. The $(t+1)^{\text{th}}$ draw of $(p_{\ell0})$ is given as
  \[
  (p_{\ell0})^{(t+1)} = \text{Truncated Normal random number} \{ (p_{\ell0})^{(t)}, \text{Var} (p_{\ell0}) \},
  \]
  where the normal distribution is truncated between 0 and 1, and Var $(p_{\ell0})$ is adaptively chosen to reduce the autocorrelation among the MCMC draws with an acceptance rate between 20% and 30%. The probability of accepting $(p_{\ell0})^{(t+1)}$ is
  \[
  \text{Min} \left\{ \frac{(L(\text{data})^{t+1} \times \text{betapdf}((p_{\ell0})^{t+1} | a, b))}{(L(\text{data})^t \times \text{betapdf}((p_{\ell0})^t | a, b))} \times \frac{\text{Tr-Normpdf}((p_{\ell0})^{t+1} | (p_{\ell0})^t, \text{Var}(p_{\ell0}))}{\text{Tr-Normpdf}((p_{\ell0})^t | (p_{\ell0})^t, \text{Var}(p_{\ell0}))}, 1 \right\}. \]

- **Generate** $C_{\ell0}, W, T, I_m, \pi_T, \pi_S, \alpha_{LAB}, \text{and } \alpha_{RPT}$

  Posterior $\{C_{\ell0}, W, T, I_m, \pi_T, \pi_S, \alpha_{LAB}, \alpha_{RPT} | p_{\ell0}, \lambda_{\ell0}, \gamma, \theta, a, b, \text{data}\}$
  \[
  \propto \prod_{i=1}^{N} L(\text{data}_i) \times \text{Prior}(C_{\ell0}, W, T, I_m, \pi_T, \pi_S, \alpha_{LAB}, \alpha_{RPT}) \tag{A3}
  \]
  Here we assume a diffuse prior for $(C_{\ell0}, W, T, I_m, \pi_T, \pi_S, \alpha_{LAB}, \alpha_{RPT})$. As (A3) does not have a closed form, the MH algorithm was used to draw from the conditional distribution of $(C_{\ell0}, W, T, I_m, \pi_T, \pi_S, \alpha_{LAB}, \alpha_{RPT})$. The $(t+1)^{\text{th}}$ draw of $(C_{\ell0}, W, T, I_m, \pi_T, \pi_S, \alpha_{LAB}, \alpha_{RPT})$ is given as
  \[
  (C_{\ell0}, W, T, I_m, \pi_T, \pi_S, \alpha_{LAB}, \alpha_{RPT})^{(t+1)} = \text{Normal random number} \{ (C_{\ell0}, W, T, I_m, \pi_T, \pi_S, \alpha_{LAB}, \alpha_{RPT})^{(t)}, \text{Var} (C_{\ell0}, W, T, I_m, \pi_T, \pi_S, \alpha_{LAB}, \alpha_{RPT}) \},
  \]
where \( \text{Var}(C_0, W, T, I_m, \pi_T, \pi_S, \alpha_{LAB}, \alpha_{RPT}) \) is adaptively chosen to reduce the autocorrelation among the MCMC draws with an acceptance rate between 20% and 30%. The probability of accepting \((C_0, W, T, I_m, \pi_T, \pi_S, \alpha_{LAB}, \alpha_{RPT})^{(t+1)}\) is

\[
\min \left\{ \frac{\prod_{i=1}^{N} L_i^{(data)}^{t+1}}{\prod_{i=1}^{N} L_i^{(data)^t}}, 1 \right\}
\]

- **Generate \((\gamma, \theta)\)**

Posterior \{\((\gamma, \theta) \mid \lambda_0\) \(\propto \prod_{i=1}^{N} \text{gammapdf}(\lambda_0 | y, \theta) \times \text{Prior}(y, \theta)\) \}

Here we assume a diffuse prior for \((\gamma, \theta)\). As \((A4)\) does not have a closed form, the MH algorithm was used to draw from the conditional distribution of \((\gamma, \theta)\). The \((t+1)^{\text{th}}\) draw of \((\gamma, \theta)\) is given as \((\gamma, \theta)^{(t+1)} = \text{Truncated Normal random number \{\((\gamma, \theta)^{(t)}\), \text{Var}(\gamma, \theta)\}}\), where the normal distribution is truncated below at 0, and \(\text{Var}(\gamma, \theta)\) is adaptively chosen to reduce the autocorrelation among the MCMC draws with an acceptance rate between 20% and 30%. The probability of accepting \((\gamma, \theta)^{(t+1)}\) is

\[
\min \left\{ \frac{\prod_{i=1}^{N} \text{gammapdf}(\lambda_0 | y, \theta)^{t+1}}{\prod_{i=1}^{N} \text{gammapdf}(\lambda_0 | y, \theta)^t} \times \frac{\text{Tr-Normpdf([(\gamma, \theta)^{t+1} | (\gamma, \theta)^t, \text{Var}(\gamma, \theta)])}{\text{Tr-Normpdf([(\gamma, \theta)^{t} | (\gamma, \theta)^t, \text{Var}(\gamma, \theta)]}, 1 \right\}
\]

- **Generate \((a, b)\)**

Posterior \{\((a, b) \mid p_0\) \(\propto \prod_{i=1}^{N} \text{betapdf}(p_0 | a, b) \times \text{Prior}(a, b)\) \}

Here we assume a diffuse prior for \((a, b)\). As \((A5)\) does not have a closed form, MH algorithm was used to draw from the conditional distribution of \((a, b)\). The \((t+1)^{\text{th}}\) draw of \((a, b)\) is given as \((a, b)^{(t+1)} = \text{Truncated Normal random number \{\((a, b)^{(t)}\), \text{Var}(a, b)\}}\), where the normal distribution is truncated below at 0, and \(\text{Var}(a, b)\) is adaptively chosen to reduce the autocorrelation among the MCMC draws with an acceptance rate between 20% and 30%. The probability of accepting \((a, b)^{(t+1)}\) is

\[
\min \left\{ \frac{\prod_{i=1}^{N} \text{betapdf}(p_0 | a, b)^{t+1}}{\prod_{i=1}^{N} \text{betapdf}(p_0 | a, b)^t} \times \frac{\text{Tr-Normpdf([(a, b)^{t+1} | (a, b)^t, \text{Var}(a, b)])}{\text{Tr-Normpdf([(a, b)^{t} | (a, b)^t, \text{Var}(a, b)])}, 1 \right\}
\]
Appendix: Parameter Recovery

In this section, we report the results of our simulation study in which we simulate data from the model for different sets of parameter values, and then estimate the model on the simulated data to test recovery of parameter values. In these simulations, we only test parameter recovery for the transactional information stock part of the model. We do not consider seasonal effects on information stock and claim characteristics in the simulation as these are more “regression like” components of the model and parameter recovery for them is expected to be good. Therefore, we consider different values of the parameters $C, W, T, \gamma, \theta, a, b$ and $\pi_T$.

We tested many different cases, but here, due to space limitations, we focus on seven sets of parameter values in the neighborhood of the parameter values that we obtain from our estimation using the health insurance data. The parameter values for the seven cases are presented in Table A1. The values in Case 0 are rounded-off values from the estimation results on the health insurance data; we include this case to ensure that the model works well for parameter values that we obtain for our focal dataset. In Cases 1 to 6, we use parameter values of comparable orders of magnitude, but introduce systematic variations to cover a variety of scenarios, explained as follows. First consider the parameters $C, W$ and $T$. In Cases 1, 2 and 3, we assume that information gain from a telephone call is higher than the information need from a claim, which is higher than the information gain from a web visit. This matches the ordering of information magnitudes obtained for our data. In Cases 4, 5 and 6, however, we flip the information magnitudes of a web visit and a claim, i.e., we assume that information gain from a telephone call is higher than the information gain from a web visit, which is higher than the information need from a claim. This allows us to test cases in which the web is an effective channel for resolving information needs from claims, though not as effective as the telephone channel. Now consider the parameters $\gamma, \theta, a$ and $b$. In Cases 1 and 4, we assume a high-heterogeneity (i.e., with a large coefficient of variation) gamma distribution for query propensity and a high-heterogeneity (i.e., bimodal and polarized) beta distribution for web-choice propensity. Relative to this, in Cases 2 and 5, we change the query propensity distribution to one with lesser heterogeneity (with a smaller coefficient of variation), and in Cases 3 and 6, we change the beta distribution for web-choice propensity to one with lesser heterogeneity (with a single interior mode). We maintain the same value of $\pi_T$ for all cases, assuming throughout that a larger information stock leads to a higher probability of a telephone call.

For each case, we simulate data for a cohort of 2500 customers for two years. We assume that each customer starts with an information stock of zero at the beginning of the two years. This information stock changes as customers file claims, which leads to information needs, which in turn induces queries by customers. To generate claims for customers, we assume that claim arrival is determined by a Poisson process for customer $i$ with the rate $\psi_i$, where $\psi_i$ is distributed across customers according to a gamma
distribution with parameters $q=1.447$ and $\delta=16.583$, such that, $f(\psi_{1}|q, \delta) = \frac{\delta q^{q}(\psi_{1})^{q-1} e^{-\psi_{1}\delta}}{\Gamma(q)}$. The values of $q$ and $\delta$ are also chosen by estimating the Poisson-gamma model on the actual claim arrival data in our health insurance dataset.

We find from our simulations that parameter recovery for the model, using the estimation procedure described in Section 5.1, is very good. For 53 of the 56 parameters recovered (8 parameters per case $\times$ 7 cases), the actual value of the parameter used to simulate the data was in the 95% credible interval obtained for the parameter from model estimation. Furthermore, the mean squared error (MSE) between the actual parameter value and the posterior mean, averaged across the seven cases, is small for every parameter, as shown in the last column of Table A1.

Next, for each case, we compute the mean absolute percentage error (MAPE) in monthly predictions of total queries, telephone queries and web queries. We then average the MAPE values across the seven cases. These averages, across the seven cases, of the MAPE values in the monthly predictions of total queries, telephone queries and web queries, are 6.07%, 13.89% and 4.46%, respectively. This shows that the accuracy in the recovery of aggregate query volumes is also high.

Overall, our simulation study indicates that parameter recovery for our model, in the neighborhood of the parameter estimates for the actual health insurance dataset, is very good. A detailed simulation study, which we leave for future work, can inform us further on this issue.

<table>
<thead>
<tr>
<th>Actual Parameter Values</th>
<th>Average MSE</th>
</tr>
</thead>
<tbody>
<tr>
<td>$C$</td>
<td>2.459E-06</td>
</tr>
<tr>
<td>$W$</td>
<td>1.265E-05</td>
</tr>
<tr>
<td>$T$</td>
<td>2.600E-05</td>
</tr>
<tr>
<td>$\gamma$</td>
<td>0.004</td>
</tr>
<tr>
<td>$\theta$</td>
<td>2.320</td>
</tr>
<tr>
<td>$a$</td>
<td>0.070</td>
</tr>
<tr>
<td>$b$</td>
<td>0.184</td>
</tr>
<tr>
<td>$\pi_T$</td>
<td>0.155</td>
</tr>
</tbody>
</table>

Table A1: Actual Parameter Values Used in the Simulations for the Seven Cases, and Average MSE in Parameter Recovery
## Appendix: Estimation Results

<table>
<thead>
<tr>
<th>Month</th>
<th>Month Index</th>
<th>$I_m$ posterior mean</th>
<th>95% cred. intv.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jul-05</td>
<td>0</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>Aug-05</td>
<td>1</td>
<td>0.123</td>
<td>[0.030, 0.189]</td>
</tr>
<tr>
<td>Sep-05</td>
<td>2</td>
<td>0.104</td>
<td>[0.029, 0.177]</td>
</tr>
<tr>
<td>Oct-05</td>
<td>3</td>
<td>0.150</td>
<td>[0.095, 0.200]</td>
</tr>
<tr>
<td>Nov-05</td>
<td>4</td>
<td>0.139</td>
<td>[0.058, 0.215]</td>
</tr>
<tr>
<td>Dec-05</td>
<td>5</td>
<td>0.121</td>
<td>[0.045, 0.183]</td>
</tr>
<tr>
<td>Jan-06</td>
<td>6</td>
<td>0.504</td>
<td>[0.432, 0.565]</td>
</tr>
<tr>
<td>Feb-06</td>
<td>7</td>
<td>0.386</td>
<td>[0.313, 0.449]</td>
</tr>
<tr>
<td>Mar-06</td>
<td>8</td>
<td>0.503</td>
<td>[0.419, 0.563]</td>
</tr>
<tr>
<td>Apr-06</td>
<td>9</td>
<td>0.262</td>
<td>[0.199, 0.341]</td>
</tr>
<tr>
<td>May-06</td>
<td>10</td>
<td>-0.019</td>
<td>[-0.092, 0.073]</td>
</tr>
<tr>
<td>Jun-06</td>
<td>11</td>
<td>-0.109</td>
<td>[-0.212, -0.018]</td>
</tr>
<tr>
<td>Jul-06</td>
<td>12</td>
<td>-0.299</td>
<td>[-0.394, -0.213]</td>
</tr>
<tr>
<td>Aug-06</td>
<td>13</td>
<td>0.265</td>
<td>[0.201, 0.321]</td>
</tr>
<tr>
<td>Sep-06</td>
<td>14</td>
<td>0.324</td>
<td>[0.222, 0.397]</td>
</tr>
<tr>
<td>Oct-06</td>
<td>15</td>
<td>0.312</td>
<td>[0.234, 0.385]</td>
</tr>
<tr>
<td>Nov-06</td>
<td>16</td>
<td>0.239</td>
<td>[0.182, 0.298]</td>
</tr>
<tr>
<td>Dec-06</td>
<td>17</td>
<td>0.115</td>
<td>[0.031, 0.188]</td>
</tr>
<tr>
<td>Jan-07</td>
<td>18</td>
<td>0.460</td>
<td>[0.374, 0.529]</td>
</tr>
<tr>
<td>Feb-07</td>
<td>19</td>
<td>0.428</td>
<td>[0.370, 0.497]</td>
</tr>
<tr>
<td>Mar-07</td>
<td>20</td>
<td>0.419</td>
<td>[0.350, 0.480]</td>
</tr>
<tr>
<td>Apr-07</td>
<td>21</td>
<td>0.270</td>
<td>[0.198, 0.335]</td>
</tr>
<tr>
<td>May-07</td>
<td>22</td>
<td>0.280</td>
<td>[0.214, 0.371]</td>
</tr>
<tr>
<td>Jun-07</td>
<td>23</td>
<td>0.311</td>
<td>[0.226, 0.363]</td>
</tr>
</tbody>
</table>

Table A2: Estimates of Seasonal Information Stock for the Information Stock Model