An empirical study of dynamic customer relationship management

Chunqing Li\textsuperscript{a,b}, Yinfeng Xu\textsuperscript{c}, Hongyi Li\textsuperscript{d,*}

\textsuperscript{a}School of Economics and Management, Tsinghua University, Beijing 100084, China
\textsuperscript{b}School of Economics and Management, Xi’an Institute of Technology, Xi’an, Shaanxi 710032, China
\textsuperscript{c}Management School, Xi’an Jiaotong University, Xi’an, Shaanxi 710049, China
\textsuperscript{d}Department of Decision Sciences and Managerial Economics, Chinese University of Hong Kong, Shatin, NT, Hong Kong

Abstract

We apply the Gönül and Shi (1998) approach to the analysis of the optimal messaging and pricing policy mix by studying the past transaction patterns between a local supermarket and its consumers. We develop a dynamic customer relationship management model and investigate the relationship between customer utility and purchasing frequency by modifying the return function of the model discussed in Gönül and Shi (1998). In particular, we extend the analysis to consider a messaging and pricing policy mix, and we use a genetic algorithm in our empirical estimation. When applied to some non-seasonal products in a local supermarket, we find that our model is suitable and far superior to the one-stage model commonly used. Our dynamic model gives the optimal marketing mix strategies in different customer states and the results show that the firm could enjoy a 22% increase in profit.

Keywords: Dynamic customer relationship management; Customer utility; Customer lifetime value; Markov-perfect equilibrium

1. Introduction

The development of information and communication technology provides many companies with a convenient environment to collect detailed data about their individual customers. These companies can take advantage of customer data to improve their marketing policies to attract more customers and thereby increase their profit. However, while many firms have invested millions in the creation of massive customer information databases, the goal of creating improved marketing policies has often proved elusive (Hoffman, 2001; King, 2001). While the trend continues of firms investing large amounts of money to collect and store customer data (Accenture, 1998), there have recently been a few high-profit firms that have reversed this current and abandoned their customer data collection efforts (Hunt, 1999). Thus, the emphasis is now on the issue of how to effectively utilize the customer databases to manage the customer relationship. It is more important to capture information about your customers than just to build up a database.

However, the potential difficulty of converting data into profits is in how to obtain relevant information from the data and customize the marketing mix policies to satisfy the consumer’s wants and needs. As pointed out in Accenture (2004), all customers are not created equal. Customer data are valuable only if it is translated into profitable interactions. Powerful analytical mathematics and statistics models are essential to identify the most profitable customers and predict future profits based on the past history of customers’ data. Modeling of customer data has become an increasingly important issue in customer relationship management (CRM). The basic idea of CRM has been embraced and the potential benefits of relationship marketing based upon individual characteristics are generally accepted (Peppers and Rogers, 1993). Over the past 10 years there has been a rapid growth of research topics on modeling the marketing mix as a function of customer relationship
status, especially taking into consideration both the customer’s utility and a firm’s profit.

Some important research considers the optimal mailing policy in direct marketing that contributes to an increasingly large share of business sales. Bult and Wansbeek (1995) construct binary choice stochastic models to determine a mailing strategy based on the response behavior of customers. DeSarbo and Rammamy (1994) consider latent class models. Neural networks are considered in Levin and Zahavi (1996). Several other papers apply a Markov decision model for direct mailing decisions, see Bitran and Mondshein (1996) and Gönül and Shi (1998), among others. In particular, Gönül and Shi (1998) were the first to propose combining the customer’s utility-maximization problem with a firm’s profit-maximization problem and thus solve the optimal timing and spacing of direct mailing of catalogs to customers.

Applying the Gönül and Shi (1998) approach, we develop a dynamic customer relationship management (DCRM) model and investigate how the marketing mix policies can be improved by using feedback from customer transaction records. We contribute to the literature in the development, estimation, and testing of an analytical DCRM model that deals with marketing mix policies. Furthermore, we model customers’ purchase behavior as a function of current and future marketing decisions in addition to the time elapsed since the last transaction and the number of continuous purchases. We also contribute in the use of a genetic algorithm when estimating the model. By solving the model we arrive at new policies that are both utility-maximizing for the customer and profit-maximizing for the firm over an infinite time horizon.

The Gönül and Shi (1998) approach gave us the initial idea of the DCRM. They present a dynamic structural model that estimates the consumer response to catalog mailings, on the assumption that consumers act as dynamic optimizers. In this context, the assumption implies that a significant part of the motivation for consumers to make purchases is to ensure that they will continue to be solicited. While this paper represents a simple type of customer management, the simple decisions to mail or not to mail are insufficient for most settings. For example, marketing tactics such as pricing and operational characteristics such as messaging are likely to have significant effects on customer retention.

This paper modifies the optimal mailing policy (OMP) model suggested by Gönül and Shi (1998). While the same hypothesis is tested using the model, we propose a dynamic customer relationship management model to illustrate techniques for converting the information contained in customer databases into meaningful knowledge and improved marketing policies. We take a mix of pricing and messaging as the main marketing tactics and measure the sensitivity of the customer’s response to these tactics. In the new model, we use the continuous purchase times to replace the accumulative purchase times so as to reduce the customer state spaces and decrease the computational complexity. We modify and verify the relationship between customer utility and continuous purchase time and modify the profit function of the firm by applying the concept of customer lifetime value (CLV) to obtain a better description of the actual CRM situation of the local supermarket under consideration. We estimate the customer discount factor as a parameter rather than taking it as a given constant which was commonly done in the literature. In this way we can take the customers’ heterogeneity into account. Finally, we utilize a genetic algorithm to estimate the model and test the hypothesis by applying the model to some non-seasonal products data from the local supermarket.

The paper is organized as follows. Section 2 presents the model of dynamic customer relationship management and the algorithm to solve the optimal marketing policy. Section 3 describes the data used in this application and the genetic algorithm used for the global optimum solution. Furthermore, the estimation results are presented. Section 4 analyzes the optimal marketing mix policy against the customer database. Finally, in Section 5 we offer our conclusions.

2. Model of dynamic customer relationship management

2.1. Research background and estimation of OMP model

The OMP model, combining a firm’s profit-maximizing problem and the customer’s utility-maximizing problem, was first presented by Gönül and Shi (1998). The object of their research was the direct mail order industry. In the OMP model, the key factors affecting the optimal mail policy were investigated by maximizing customer utility and the direct-mail company’s profit. This model utilized the estimable structural dynamic programming (ESDP) technology to model customer utility and the firm’s profit. They considered the exchange between the customer and the firm as a stochastic game process and an equilibrium solution was obtained. The main contribution of the model is that the authors built, estimated, and tested one model that manages the mail policy. The components the model considered were the recency of the last purchase, the accumulative purchase time that describes the customer’s purchasing behavior, and the mail policy. Using the model, a firm can get the optimal mail policy to maximize customer utility and its own profit in an infinite time horizon.

The basic assumption in the OMP model is that both the firm and the customers consider not only the current benefit but also the future benefit of a decision. The
information between the customers and the firm is completely symmetrical and the stochastic errors (ε_{it}) are assumed to follow an iid standard normal distribution in customer utility.

The model has some problems, which are detailed in the following. (1) There are no detailed algorithms for the global optimal solution in the estimation process. (2) The customer state space would be very large if the customer makes purchases frequently, which will increase the complexity of the computation. (3) The mail policy is not sufficient for customer relationship management. (4) The constant customer discount factor cannot well reflect the consumers’ heterogeneity.

2.2. Ideology of DCRM modeling

In general, the philosophies and tactics that emphasize the importance of individual customers are labeled as CRM. In the DCRM model, the term “dynamic” has three meanings. The first is that both the customers and the firm in the model are “forward looking” when they make decisions. That is to say, the firm and the customers pay attention to not only the current benefit but also the future benefit of their behavior. The second is that we consider both the firm’s and the customers' benefit when building the model, that is, we investigate the key determinants of the optimal marketing mix policies in a dynamic environment where customers maximize utility and the firm maximizes profit. The third meaning of the term “dynamic” is that we focus on the relationship rather than the transaction. The customers are the source of repeated transaction; therefore, the value of the customers exceeds the value of their current transactions. Customers should be considered as an important equity, which brings the firm profits when the customers make repeated purchases. With this understanding, the costs the firm has to bear to retain the customers may be considered as an investment. Similarly, the related marketing policies on the existing customers should be considered as the management of existing assets. The OMP model presents the theoretical foundation and the rudiments of the DCRM.

In each period, the firm selects the marketing mix policies (or controls) to offer the customers in each state and the consumers decide whether or not to buy the product in a given period. This builds the framework for a multistage repeated game. The customers’ decisions, influenced by the firm’s marketing actions, define the transitions between the states. From the perspective of the firm, the consumers’ decisions and the system’s transition probabilities are random variables. Under the assumption that the customers’ purchasing decisions are a function of the current customer state and the firm’s policies, the sequence of customer states forms a Markov chain process. Since the firm has the ability to (partially) control the evolution of the Markov chain process through the selection of control variables, customer management may be treated as a Markov decision process (Lewis, 2001).

2.3. The customer behavior model of DCRM

2.3.1. The objective function and response variable of the customer

The objective function of the customer is the same as that in the OMP model in Gönül and Shi (1998). It is to maximize long-term utility while making purchasing decisions in each period over an infinite time horizon, i.e.,

$$\max \left( E \left( \sum_{t=1}^{\infty} \delta^{-t} u_{it} d_{it} \right) \right),$$

where $\delta_t$ is the customer’s discount factor, which is the inverse of $(1 + \text{discount rate})$, $u_{it}$ is the utility of the $i$th customer making a purchase at period $t$, and $d_{it}$ is a binary variable denoting the customer response, which is defined as follows:

$$d_{it} = \begin{cases} 1 & \text{if individual } i \text{ makes a purchase at period } t, \\ 0 & \text{otherwise}. \end{cases}$$

The customer’s response variable is mutually exclusive and sole. In fact, the customer not only decides whether to make a purchase or not, but also decides the amount and the brand of the purchase. To simplify our calculation, we assume that the customer only purchases one commodity when he/she makes a purchase. Moreover, the brand of the commodity purchased holds the majority of the market share. This allows us to avoid questions about the amount of purchase and the choice of brand; however, the conclusions should remain unchanged. The customer’s state variable and its evolution in our model, described in Sections 2.3.2 and 2.3.3, are different from those in the OMP model in Gönül and Shi (1998).

2.3.2. Customer’s state space and its evolution

Recency, frequency, and monetary value are the fundamental factors to describe a customer’s purchase behavior, according to the famous RFM model in Bitran and Mondschein (1996). Recency (R) stands for the time elapsed since the last purchase. Frequency (F) refers to the number of purchases in the past or the proportion of purchases over a period of time. Monetary (M) value is the amount spent so far or the

---

1. Viaene et al. (2001) extend the standard RFM modeling by adding other non-RFM predictors and including alternative operationalizations of the RFM variables. Due to computational complexity, we do not consider these in our modeling.
average amount spent per purchase so far. The model is derived from the direct marketing industry (Dwyer, 1989) and has recently been extended to other industries (Lewis, 2001). As in this paper we only consider whether a customer makes a purchase or not, we use recency (r), the time elapsed since the last purchase, and frequency (f), the number of continuous purchases in the past or the proportion of purchases over a period of time. In our model, there are some differences to the previous definitions of the RFM model. The two variables are defined as the customer’s state variables. We define frequency here as the number of continuous purchases so as to decrease the customer’s state space and thus reduce the computational complexity.

Both the recency and frequency for an individual i evolve according to a Markov transition process. Therefore we have

$$r_{i,t+1} = \begin{cases} 0 & \text{if } d_{it} = 1, \\ r_{it} + 1 & \text{if } d_{it} = 0, \end{cases} \quad (3a)$$

and

$$f_{i,t+1} = \begin{cases} 1 & \text{if } d_{it} = 1, \\ 0 & \text{if } d_{it} = 0. \end{cases} \quad (3b)$$

Eq. (3a) indicates that recency reduces to 0 if $d_{it} = 1$, and increases by 1 if $d_{it} = 0$, while Eq. (3b) states that frequency reduces to 0 if $d_{it} = 0$, and increases by 1 if $d_{it} = 1$. Here we modify the meaning of frequency from accumulative purchase times to continuous purchase times. The advantage of doing this is that it not only reduces the customer state spaces which ultimately decreases the computational complexity, but also reflects the common practice in retailing industry. We give a comparison of the state spaces and the time variables between DCRM and OMP models in Table 1.

For customers who make initial purchases in a supermarket, we set $r = 0$ and $f = 1$. For the customer’s state space, $S_{it} = \{r_{it}, f_{it}\}$, please refer to Gönül and Shi (1998).

### 2.3.3 Customer’s utility function and valuation function

In accordance with the customer’s state (r and f) and the firm’s marketing mix policy variables, according to the characteristics of the supermarket, the firm’s marketing mix policies mainly include messaging policy and price policy as described in Eqs. (8) and (9); we model the customer’s purchase utility as follows:

$$u_{it} = \pi + \beta_1 mm_{it} + \beta_2 p_{it} + \beta_3 r_{it} + \beta_4 f_{it}^2 + \beta_5 \ln(f_{it} + 1) + \epsilon_{it} = \bar{u}_{it} + \epsilon_{it}, \quad (4)$$

where $\bar{u}_{it}$ represents the deterministic component of the utility function and $\epsilon_{it}$ represents the random component.

We modify the intrinsic utility function in the OMP model by adding a price variable ($p_{it}$) to the utility function while changing the mail catalog into messaging ($mm_{it}$), which mainly includes phone and e-mail messages. The relationship between the customer’s utility and frequency is specified as a logarithmic relationship, which turns out to be significant in our empirical study, as discussed later on. Similar to Gönül and Shi (1998), we normalize the utility of not buying in a single period to zero.

Thus, the customer’s valuation function in period t is

$$V_{it}(S_{it}) = \begin{cases} \bar{u}_{it} + \delta t \bar{EV}_{it+1} & \text{if } d_{it} = 1, \\ \delta t \bar{EV}_{it+1}(S_{it+1}|d_{it} = 0) & \text{if } d_{it} = 0. \end{cases} \quad (5)$$

The expected valuation function is

$$EV_{it}(S_{it}) = \text{Prob}_0(d_{it} = 1|S_{it}, mm_{it}, p_{it}) \times [\bar{u}_{it} + \delta t \bar{EV}_{it+1}(S_{it+1}|d_{it} = 1)] + \text{Prob}_0(d_{it} = 0|S_{it}, mm_{it}, p_{it}) \times \delta t \bar{EV}_{it+1}(S_{it+1}|d_{it} = 0) + \Phi [\delta t \bar{EV}_{it+1}(S_{it+1}|d_{it} = 1) - \bar{u}_{it}], \quad (6)$$

where $\Phi$ is the standard normal cumulative distribution function and $\delta t$ is the customer discount factor. The state variables evolve according to Eq. (3). The customer purchase probability is simply that the probability of the long-term utility of purchasing is more than the long-term utility of not purchasing. Thus, the customer purchase probability is

$$\text{Prob}_0(d_{it} = 1|S_{it}, mm_{it}, p_{it}) = \text{Prob}_0(\bar{u}_{it} + \delta t \bar{EV}_{it+1}(S_{it+1}|d_{it} = 1) + \epsilon_{it} > \delta t \bar{EV}_{it+1}(S_{it+1}|d_{it} = 0)) = \Phi [\bar{u}_{it} + \delta t \bar{EV}_{it+1}(S_{it+1}|d_{it} = 1) - \bar{EV}_{it+1}(S_{it+1}|d_{it} = 0)]. \quad (7)$$

There are detailed estimation steps to solve Eqs. (5)–(7) in Gönül and Shi (1998). The expected valuation

### Table 1

<table>
<thead>
<tr>
<th>Model</th>
<th>Expression of state spaces</th>
<th>t</th>
</tr>
</thead>
<tbody>
<tr>
<td>DCRM</td>
<td>$2(t - 1)$ and $t \neq 1$</td>
<td>1</td>
</tr>
<tr>
<td>OMP</td>
<td>$(t - 1)t/2 + 2$</td>
<td>1</td>
</tr>
</tbody>
</table>
function is evaluated recursively using the algorithm presented in Section 2.5 below.

2.4. Firm’s behavior in the DCRM model

The firm’s objective is to maximize the profit from individual customer purchases. In other words, the firm maximizes the customer lifetime value (CLV), defined as the expected discounted contribution produced over the lifetime for a given customer (see Kotler and Armstrong, 2001). For an individual customer, the customer equity is equivalent to the CLV. The CLV has long been advocated as the appropriate criterion for evaluating direct-marketing programs for customer acquisition (Dwyer, 1989). It is becoming an increasingly popular measure for firms in more traditional marketing environments (Blattberg and Deighton, 1996; Berger and Nasr, 1998; Blattberg et al., 2001; Reinartz and Kumar, 2000, 2003). The key insight behind these concepts is that a firm’s customer base may be viewed as a portfolio of assets that produce revenue streams over time that exceed the company’s costs. Specifying CLV as the firm’s objective function makes it clear that customer management should focus on how marketing actions influence the expected stream of future revenue over costs rather than only on the current transaction.

2.4.1. Firm’s marketing policy space

It is a well-known fact that some prices are more often used in retailing than others (Aalto-Setälä et al., 2004). Furthermore, messaging is the main method of communication in 4Cs (customer, cost, convenient, and communication). Accordingly, out of all of the marketing mix policies, we only consider the message policy and the price policy to best describe the business characteristics of the local supermarket considered in our empirical study. Thus, the marketing mix policy space of the firm for the $i$th customer in period $t$ is $D_{it} = \{mms_{it}, p_{it}\}$. Here $mms_{it}$ stands for the messaging individual $i$ in period $t$ via telephone, e-mail, mobile phone (messages), and so on. It is a binary variable defined as

$$\text{mms}_{it} = \begin{cases} 1 & \text{if messages for individual } i \text{ in period } t, \\ 0 & \text{otherwise.} \end{cases}$$

(8)

The other policy variable, $p_{it}$, stands for the firm’s price policy for individual $i$ in period $t$. It is defined as the percentage price change compared to the primary price. That is,

$$p_{it} = \frac{P_{it} - P_0}{P_0},$$

(9)

where $P_{it}$ stands for the price for $i$ in period $t$ and $P_0$ stands for the primary price. In the actual computation, $p_{it}$ should be a discrete variable to facilitate the estimation.

2.4.2. The firm’s profit model

To simplify the model, we assume that customers can only purchase one commodity at each purchase and that the stock cost of the commodity is uniform in all periods. Then the profit of the firm from individual $i$ in state $S_{it}$ in the current period $t$ is

$$\pi_{it}(S_{it}, mms_{it}, p_{it}) = R(p_{it}) \text{Prob}_{it}(d_{it} = 1|S_{it}, mms_{it}, p_{it})$$

$$- c \times mms_{it},$$

(10)

where $c$ is the unit messaging cost and $R(p_{it})$ stands for the net revenue from individual $i$ when the price policy is $p_{it}$. That is,

$$R(p_{it}) = P_{it} - C = P_0(p_{it} + r_0),$$

(11)

where $r_0$ stands for the gross profit rate for the primary price $P_0$. That is,

$$r_0 = \frac{P_0 - C}{P_0},$$

(12)

where $C$ stands for the stock cost of the commodity. Then, we define $CLV_{it}(S_{it})$ as the maximum expected profit that customer $i$ contributes in state $S_{it}$ in the time interval $(t, \infty)$:

$$CLV_{it}(S_{it}) = \max_{mms_{it}, p_{it}} \{\pi_{it}(S_{it}, mms_{it}, p_{it})$$

$$+ \delta_{it} \left[ \text{Prob}_{it}(d_{it} = 1|S_{it}, mms_{it}, p_{it}) \times CLV_{i,t+1}(S_{i,t+1}|d_{it} = 1) \right.$$

$$+ \text{Prob}_{it}(d_{it} = 0|S_{it}, mms_{it}, p_{it}) \times CLV_{i,t+1}(S_{i,t+1}|d_{it} = 0) \} \},$$

(13)

where $\delta_{it}$ is the firm’s discount factor (which we set at 0.991 to yield a 10% annual discount rate). Here $r$ and $f$ evolve according to Eq. (3).

2.5. Solving the firm’s optimal marketing mix policy

We consider a stochastic game process and the ESDP to solve the firm’s optimal marketing mix policy. As in our model, only the customer’s state space and the firm’s policy space are different from those in the original OMP model discussed in Gönül and Shi (1998), the conditions for the existence and the uniqueness of the solution are satisfied. The solution is computed by Algorithm 1 discussed below.

Algorithm 1

1. **Initialization.** Let the profit function $CLV_{i,t+1}(S_{i,t+1}) = 0$ and the expected value function $EV_{i,t+1}(S_{i,t+1}) = 0$ for all state variables. Set a tolerance level $\eta > 0$.
Step 1: Use Eq. (7) to compute \( \text{Prob}_{it} \) \( (d_{it} = 1|S_i, mms_{it}, p_{it}) \), the response probability of customer \( i \) in state \( S_i(r_{it}, f_{it}) \), for the firm’s marketing mix policy \( D_i^*(mms_{it}, p_{it}) \).

Step 2: Use Eq. (5) to compute \( V_i(S_i) \), the value function of customer \( i \) in state \( S_i(r_{it}, f_{it}) \).

Step 3: Use Eq. (13) to compute the maximal expected profit \( CLV_i(S_i) \) from customer \( i \) and the corresponding optimal marketing mix policy \( D_i^*(S_i) \).

Step 4: Use Eq. (8) to compute the expected future value of customer \( i \) in state \( S_i \) by employing the optimal marketing mix policy.

Step 5: Stopping criterion. Let \( d_1 = EV_{i,t+1} - EV_{it} \) and \( d_2 = CLV_{i,t+1} - CLV_{it} \). If \( d_1^*d_2 < \eta \), then stop. Otherwise let \( EV_{i,t+1} = EV_{it} \) and \( CLV_{i,t+1} = CLV_{it} \) and go to Step 1.

Thus, we obtain the optimal marketing mix policy that is both profit maximizing for the firm and utility maximizing for the individual customer over an infinite time horizon. For more details of the algorithm, please see Göñül and Shi (1998).

3. An empirical application estimation of the DCRM model

We use data from one local supermarket in Xi’an, China (the name of the supermarket is kept confidential at the request of the company). The supermarket specializes in non-perishable groceries and daily necessities. The data include all the customers’ transactions for the supermarket’s first 30 months of operations. In this time period the firm received approximately 120,000 orders from 50,000 plus customers. For each order, a record is kept of the customer’s identifying information, such as customer number, address, purchase time, posted price, and consumption spending.

The sample log-likelihood function is therefore \( LL = \sum_{i=1}^{I} \sum_{j=1}^{n_{ij}} L_{ij} \), where \( I \) is the number of customers and the interval \( [b_i, B_i] \) is the interval during which individual \( i \) is observed. We expect that the customers can acquire positive utility through reading the message from the firm. Therefore, the expected coefficient of messaging is positive. However, customer utility will decrease with increasing prices. The expected coefficient of the price is therefore negative. There exists a U-shaped relationship between \( r \) and customer utility, as discussed in Göñül and Shi (1998).

3.1. Data

Our data set includes all the purchase history of the supermarket’s individual customers from June 2000 to December 2002. The commodity is non-seasonal and is consumable all year round. In our sample, we chose customers whose first purchases were made in May 2000. The number of customers in the sample is 580. On average, the number of messages received yearly for every customer is about seven, with the cost of each message being 0.2 RMB yuan. There are three price levels in the primary marketing policy. They are 23 (the stock cost) and 24 and 25 (the original price) yuan, with the price fluctuation ratios being –0.08, –0.04, and 0, respectively. Because the unitary stock cost of the commodity commonly fluctuates above the level of 23 yuan, we think that the unitary stock cost of the commodity is 23 yuan. Then rate0 in Eq. (12), namely the gross profit rate for the primary price of 25 yuan, is 0.08. The gross profit rate for the primary price is 0.08. We summarize the basic characteristics of the data in Table 2.

3.2. The likelihood function and the estimation algorithm

The log-likelihood function for customer \( i \) in period \( t \), which consists of the products of the response and non-response probabilities of the customers over the observation period, is given by

\[
L_{it} = d_i \ln(\text{Prob}_{it}(d_i = 1|S_i, mms_{it}, p_{it})) + (1 - d_i) \times \ln(\text{Prob}_{it}(d_i = 0|S_i, mms_{it}, p_{it})).
\]

\[ (14) \]

The sample log-likelihood function is therefore \( LL = \sum_{i=1}^{I} \sum_{j=1}^{n_{ij}} L_{ij} \), where \( I \) is the number of customers and the interval \( [b_i, B_i] \) is the interval during which individual \( i \) is observed. We expect that the customers can acquire positive utility through reading the message from the firm. Therefore, the expected coefficient of messaging is positive. However, customer utility will decrease with increasing prices. The expected coefficient of the price is therefore negative. There exists a U-shaped relationship between \( r \) and customer utility, as discussed in Göñül and Shi (1998).

### Table 2: Summary statistics

<table>
<thead>
<tr>
<th>Sample characteristics</th>
<th>Mean</th>
<th>Std. dev.</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of purchases per customer</td>
<td>4.2086</td>
<td>2.3250</td>
<td>1</td>
<td>14</td>
</tr>
<tr>
<td>Number of messages received per year per customer</td>
<td>7.1086</td>
<td>2.1968</td>
<td>1</td>
<td>14</td>
</tr>
<tr>
<td>Number of prices of 23 yuan encountered per customer</td>
<td>13.4534</td>
<td>5.2029</td>
<td>0</td>
<td>25</td>
</tr>
<tr>
<td>Number of prices of 24 yuan encountered per customer</td>
<td>16.0586</td>
<td>4.9666</td>
<td>5</td>
<td>30</td>
</tr>
<tr>
<td>Number of prices of 25 yuan encountered per customer</td>
<td>0.4879</td>
<td>1.1419</td>
<td>0</td>
<td>10</td>
</tr>
<tr>
<td>Number of months between purchases</td>
<td>5.1353</td>
<td>5.6338</td>
<td>1</td>
<td>29</td>
</tr>
<tr>
<td>Number of continuous purchases per customer</td>
<td>1.2328</td>
<td>1.4970</td>
<td>0</td>
<td>11</td>
</tr>
<tr>
<td>( r ) Value when making purchase</td>
<td>4.3245</td>
<td>5.6324</td>
<td>0</td>
<td>29</td>
</tr>
<tr>
<td>( f ) Value when making purchase</td>
<td>0.5920</td>
<td>1.1318</td>
<td>0</td>
<td>11</td>
</tr>
</tbody>
</table>

Note: Number of customers in the sample is 580. Total number of months observed is 17,400.
Shi (1998). There also exists a logarithmic relationship between $f$ and customer utility. The expected coefficient of $r$ is negative; the expected coefficient of $r^2$ is positive; and, finally, the expected coefficient of $f$ is positive.

The models in our paper represent a combination of dynamic programming and discrete choice econometrics. Dynamic programming evaluates the consequences of each alternative at each decision point and choice modeling techniques assess the likelihood of the observed choices and search for parameter values that best fit the data. The standard approach to implementing these types of models is to construct an algorithm including the dynamic programming procedure to maximize sample log-likelihood by constantly adjusting the parameter values in the utility model.

In the estimation of the maximum likelihood function, we first solve the stochastic dynamic programming problem using the backward recursion algorithm conditional on a set of parameter values. Then we evaluate the sample log-likelihood function. This is done in a two-step iteration with different parameter values to find out the sample log-likelihood until convergence. Because the objective function is similar to multimodal distribution and has complicated nested relations, it is not applicable to use the gradient analytic method to search global optima due to the tendency to get stuck into local optima and the difficulties of calculating gradients. To reach the global optima, we choose genetic algorithms. Genetic algorithms are search algorithms based on natural genetics and selection, which combine the concept of survival of the fittest with a structured interchange and aleatory changes by means of crossover and mutation operations. The greatest advantage of genetic algorithms as opposed to traditional search methods, such as the gradient method, is the ability to avoid local optima (Streifel et al., 1999). Another advantage is its utility in real-time applications. Despite not providing the optima to the problem, it provides a solution that is almost better in a shorter time (D’ummler, 1999). Genetic algorithms are especially suitable for solving and optimizing problems, not only because that they do not get trapped in local optima when their parameters are set appropriately, but also because they show quicker convergence than other stochastic optimization algorithms, especially for high-dimensional problems (Michalewicz, 1996). Genetic algorithms naturally use binary strings to code solutions, which allow an entire set of multi-objective optima to be evolved simultaneously ( Fonseca and Fleming, 1995). The idea of combining genetic or evolutionary algorithms with economic problems (Arifovic, 1994; Price, 1997; Cooper, 2000; Varetto 1998) and econometrics problems (Dorsey and Mayer, 1995; Venkatesan et al., 2004) is not new. However, the application of genetic algorithms to marketing sciences is relatively new. Using the following algorithms, we obtain a global solution quickly.

### The genetic algorithm

**Step 0:** Randomly produce the parameter set of the customer utility equation. Then, transform them into binary code to form chromosomes in string format; each chromosome represents an individual. The process is repeated 50 times, that is, the population size is 50. Set a tolerance level $\eta > 0$ (in this case it is set to 0.01).

**Step 1:** Use Eq. (14) to compute the sample log-likelihood $LL_i$ for every individual in the parent group and find the minimum sample log-likelihood $LL_{min}$. Then, define the fitness function as $f_i : LL_i - LL_{min}$.

**Step 2:** According to the principle of “Better (fitter) individuals have a higher chance of survival,” the probability $p_i$ for the individual to be chosen will be higher if the fitness of the individual is better. So we define $p_i$ as $f_i / \sum_{i=1}^{N} f_i$. According to the “elite principle”, we will always retain two of the best individuals (the two individuals whose fitness functions $f_i$ are the highest). Then, we randomly choose two individuals in the parent group, which excludes the two best individuals, as the parent according to probability $p_i$.

**Step 3:** For the parent, we randomly choose the crossover points in both chromosomes and cut them off at the crossover points. The parent exchanges the parts of the chromosomes following on from the crossover point to get two new chromosomes (named “offspring”).

**Step 4:** A mutation rate (probability that a certain gene mutates) of 1% is applied. A chromosome undergoes mutation, in which one of its attributes changes from 0 to 1 or from 1 to 0. We randomly choose some individuals according to the mutation rate and then randomly choose one part of them. We change the value to 0 if the primary value is 1, or change the value to 1 if the primary value is 0.

**Step 5:** The stopping condition. If $\sum_{i=1}^{N} f_i \leq \eta$, or if the iteration of the process exceeds 5000, then stop. Otherwise go to Step 1.

In our empirical analysis, the number of iterations is 1248.

#### 3.3. Estimation results

The dynamical model was presented in the above sections. Now, for the purposes of comparison, we consider a single-period model where the customers do not value future utility at all ($\delta_c = 0$). The two models...
are different in that the values of the customer discount factor $d_c$ can be different. We rank the models using the Akaike information criterion (AIC) and the Bayesian information criterion (BIC) (Gönül and Shi, 1998; Lewis, 2001). In contrast to previous research (say, Gönül and Shi, 1998), we obtain the value of the discount factor $d_c$ by estimating it as a parameter, rather than by treating it as a given constant value. We apply three approaches to determine the customer discount factor. The customer’s discount factor ($d_c$) plays a key role in the customer choice models. It directly influences the firm’s optimal marketing mixed strategies. Most literature (Gönül and Shi, 1998; Lewis, 2001; Wolpin, 1984; Gönül, 1989; Hotz and Miller, 1993) considers the same method in the determination of the discount factor. They use a constant customer’s discount factor, e.g., the monthly customer’s discount factor will be 0.991 and the weekly customer’s discount factor will be 0.998. In this way the annual discount rate will be 10%. Such a method simplifies the model, but it is not perfect because it cannot reflect the customer heterogeneity. Gönül and Shi (1998) suggest that the customer’s discount factor should be estimated as a parameter. However, it might cause a problem of non-convergence. We estimate the factor as a parameter and do not encounter the possible non-convergence problem. We find that the dynamic model best reflects the characteristics when the value of the customer discount factor $d_c$ is 0.7. The estimation results are presented in Table 3.

According to both the AIC and the BIC indices, the dynamic model performs better than the single-period model. The parameter estimation of the dynamic model indicates the following implications. Messaging generates a small but significant positive effect on customer response probability. Similarly, an increase in price decreases the customer’s response probability, while a decrease in price increases the customer’s response probability. Therefore, the impact of recency is a U-shaped curve (see Fig. 1), while the impact of frequency is roughly logarithmic (see Fig. 2). Our estimation results seem to verify our hypotheses about recency and frequency. Moreover, we reject the monotone relationships for both variables in this product category. Note that in Figs. 1 and 2, the vertical axis $prob$ stands for the response probability.

### 4. The optimal marketing mix policy

Before determining the optimal marketing mix policy, we convert the continuous variable, $p_i$, into a discrete variable. We set the upper limit at 0 and the lower limit at $-0.1$. Let $p_i$ increase by 0.01 at each step from the lower limit $-0.1$ to the upper limit 0.

We emphasize here that the results discussed below are only valid for the database of the particular local supermarket that we investigated. For other cases the model may need to be modified. Since we obtain the equilibrium solution over an infinite time horizon using the algorithm presented in Section 2.5, the optimal marketing mix policy should be uniform when the recency and frequency are fixed even if the time period is different. To better explain the result, we have chosen

### Table 3: Estimation results of the single-period and the dynamic models

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Single-period model</th>
<th>Dynamic model</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta$</td>
<td>$-1.38^{***}$</td>
<td>$-1.51^{***}$</td>
</tr>
<tr>
<td></td>
<td>($-3.7902$)</td>
<td>($-4.2187$)</td>
</tr>
<tr>
<td>$\beta_m$</td>
<td>0.08</td>
<td>0.11</td>
</tr>
<tr>
<td></td>
<td>(0.6379)</td>
<td>(0.9203)</td>
</tr>
<tr>
<td>$\beta_p$</td>
<td>$-5.1$</td>
<td>$-6.1$</td>
</tr>
<tr>
<td></td>
<td>($-0.7235$)</td>
<td>($-0.8803$)</td>
</tr>
<tr>
<td>$\beta_D$</td>
<td>$-0.082^{**}$</td>
<td>$-0.088^{**}$</td>
</tr>
<tr>
<td></td>
<td>($-1.8453$)</td>
<td>($-2.0144$)</td>
</tr>
<tr>
<td>$\beta_2$</td>
<td>0.0043***</td>
<td>0.0037***</td>
</tr>
<tr>
<td></td>
<td>(3.1973)</td>
<td>(2.7986)</td>
</tr>
<tr>
<td>$\beta_f$</td>
<td>0.68***</td>
<td>0.60***</td>
</tr>
<tr>
<td></td>
<td>(3.5199)</td>
<td>(3.2551)</td>
</tr>
</tbody>
</table>

Note: Significance at the 0.01 level is denoted by ‘***’, at the 0.05 level by ‘**’, and at the 0.1 level by ‘.’. The asymptotic normal test statistics are in parentheses.
the 25th period as a typical example for discussion, because the characteristics of the result in this period are the most remarkable.

(1) We analyze the conditions that the value of $f$ is zero, that is, the customer has not made a purchase since at least the previous month. To such a customer, the optimal price policy the firm should use is presented in Fig. 3. Note that in Fig. 3, the condition that the value of recency can be either six or nine might be caused by noises. In Fig. 3, $p$ stands for the price policy. On the other hand, the optimal messaging policy has values $mms = 0$ for all $r$. Because of the low messaging sensitivity and the high price sensitivity of this group of customers, the firm does not need to send them messages. This can decrease the unnecessary cost of messaging. Instead, the firm should use the corresponding price policy. When the time for not buying is only one month ($r = 1$), the firm does not need to worry about these customers, because the reason for not buying may be that the commodity bought in the previous period is still sufficient or may be due to some random reason. These customers may come to make purchases at any moment in the next period. Therefore, the firm does not need to offer discounts to them. When the time of not buying is between 2 and 15 months ($2 \leq r \leq 15$), the probability that the customers will make a purchase is low, as shown in Fig. 1. In this case, the customers probably have many other companies as alternatives. They would rather compare prices than services when the qualities are the same. Thus, the firm should consider a discount price policy to attract them. When the time for not buying exceeds 15 months ($r > 15$), the probability that the customers will make a purchase is high, as shown in Fig. 1. These customers may also come to make purchases at any moment. Therefore the firm does not need to offer them any discount.

(2) We analyze the situation when the value of $r$ is zero, that is, the customer has made a purchase in at least the previous month. To such a customer, the optimal messaging policy the firm should consider is presented in Fig. 4. On the other hand, the optimal price policy has values $p = 0$ for all $f$. Because of the low price sensitivity and the high messaging sensitivity of this group of customers, the firm does not need to offer them discounted prices to achieve greater profits. The firm can just use the corresponding messaging policy. When the time for continuous buying is between one and 7 months ($1 \leq f \leq 7$), the firm does not need to send these customers messages. One of the possible reasons for their continuous purchases could be that it is convenient to obtain the commodity in this local supermarket. They are not necessarily loyal to the product. Therefore these customers do not demand quality of service. When the time for continuous buying is between 8 and 15 months ($8 \leq f \leq 15$), these customers regard themselves as the loyal customers of the firm or the product. They may complain if the quality of service does not satisfy them. This could damage the relationship between the customers and the firm. At this moment, the firm should treat these customers as important loyal customers. These customers would rather value quality of service than prices. This is similar to the findings in Reichheld and Teal (1996). The firm should consider the use of the messaging policy but not the discount price policy. In this way the firm can maintain or even strengthen, step by step, the loyal relationship between the firm and the customers through the messaging policy to provide high-quality services to these customers. With the customers’ continuous buying, the loyal relationship between the firm and the customers would be strengthened. The customers are now used to making purchases from the firm and the firm may even save the cost of messaging when the time for continuous buying exceeds fifteen ($f > 15$) (see Fig. 4). Thus, we find that the relationship between the firm and the customers starts to be a close and loyal one when the time for continuous buying is eight. The firm should take this opportunity to lower the cost of messaging, while at the same time strengthening the loyal relationship to realize more profit for the firm. Note that in Fig. 4, the situation where the value of frequency is 18 may be caused by noises. Also note that in Fig. 4 $mms$ stands for the messaging policy.

After comparing the profitability of the optimal marketing mix policy and the primary marketing mix policy, we find that the firm’s profit would be 3520 yuan in period 30 after use of the optimal marketing mix policy. This amount exceeds the primary profit of 2876 yuan.

![Fig. 3. Optimal price policy for recency.](image1)

![Fig. 4. Optimal messaging policy for frequency.](image2)
yuan from the primary marketing policy. There is an increase in profit of 22%.

5. Conclusions

This paper constructs a DCRM model for a local supermarket. We apply the Göñül and Shi (1998) approach in the analysis of the transaction patterns between the local supermarket and its consumers. We extend the analysis to consider a marketing mix policy of messaging and pricing. The results show that the firm’s marketing mix policy should be flexible according to the customer’s individual recency and frequency values. Customers whose recency values are not zero value prices rather than service. Therefore, the firm should consider the price policy rather than sending messages to them. The firm does not need to provide discount to customers when the time elapsed is either too short or too long. Instead, the firm should consider the discount policy when the recency is moderately long.

For customers who make continuous purchases, the firm should send messages to them rather than providing price discounts. However, the firm does need to send messages when the continuous purchasing times are either too short or too long. When the frequency is between the two, the firm should consider the discount policy to promote the relationship between the firm and the customer to a loyal stage. The results also show that the profit from the optimal marketing mix policy exceeds the profit from the primary marketing mix policy, which has important implications for the profit management of the firm.

Acknowledgments

Chunqing Li and Yinfeng Xu would like to acknowledge the support of the NSF of China (10371094, 70121001) and the Science Foundation of Shaanxi Province (02G11, 03G07). Hongyi Li would like to acknowledge the financial support given by the Direct Grant, BA Faculty, of the Chinese University of Hong Kong. We would like to thank Professor Harry Timmermans, the editor, and an anonymous referee for their helpful comments. The remaining errors are, however, our own.

References


Reinartz, W., Kumar, V., 2000. On the probability of long-life customers in a noncontractual setting: an empirical investigation
and implications for marketing. Journal of Marketing October 64, 17–35.