Examining credit card consumption pattern

Yuhao Fan
(Economics Department, Washington University in St. Louis)

Abstract: In this paper, I analyze the consumer’s credit card consumption data from a commercial bank in China. To do so, I refer to the SMC model proposed by Schimittlein et al. in 1987. I am proposing an extension to the model to incorporate the credit card spending amount into the model. I use transaction frequency, recency (defined as how recent a customer has used the credit card) and spending amount data to model consumer credit card consumptions and estimate parameters with the data. I get reliable estimates from the model. I introduce a value measure that is less informative but easier to compute than traditional CLV measures, to measure consumers’ contribution to the credit card issuer (the value of credit cards). I define the value measure of a credit card as the expected spending amount next month conditional on the card still active in this month. I further regress the value measures against the characteristics information of consumers and credit card to examine which characteristics information are the best predictors of values of the credit cards. I find that characteristic information is not useful estimates of the value measure of a credit card. Rather, the on-going value of a credit card is best explained by prior consumption records.

Keywords: credit card consumption, Pareto/NBD “counting your customers” model, recency, frequency, purchase amount

1 This is my Senior Honor Thesis, and I am working with Professor Antinolfi in the Economics Department of WUSTL.
1. Introduction

Understanding consumption pattern is important for customer relationship management. This understanding allows firms to discriminate among consumers and make strategic marketing plans. For example, Kumar et al (2008) report that IBM keeps scores on their customers’ historical behaviors and use these scores to take marketing actions.

In both industry and academia, there has been great interest to develop models to predict consumption patterns. The SMC Model proposed by Schmittlein et al. (1987) is the earliest model, with a few extensions to the model proposed by other authors. Schmittelein et al. (1987) assume that purchases of each customer occur at random times and follow a Poisson process. The customers may drop out of the market randomly and the lifetime of the customers follows an Exponential distribution. Furthermore, they assume that heterogeneities of the parameters across the customers follow Gamma distributions. Under our assumptions, customers’ purchases show different patterns which are incorporated with different parameters, but these parameters are all from the same Gamma distributions and are thus connected. Although data on one single customer is limited, with a large population, we can still get accurate estimation of the parameters. Once we finish the model estimations, we can draw statistics of managerial relevance for marketing purpose, including the probability of being active at a certain point in time, expected lifetime, one-year survival rate, and expected number of transactions in a future period. The model has a number of successful applications. For example, Fader et al. (2005) study the e-commerce transactions over 78 weeks; Abe (2009) examines the shopping records for the members of a frequent shopper program at a department store in Japan for 52 weeks starting from July 1, 2000.

My data is from a commercial bank in China. It records its consumers’ credit card purchase behaviors and characteristics information of its consumers and the cards issued. In this paper, I use the model to examine consumers’ credit consumption patterns. To my knowledge, this is the first attempt to examine credit card consumption patterns with the SMC model. The model mentioned above focuses on the recency (defined as the most recent purchase of the card) and frequency of the purchases. They do not account for spending amount of each purchase. Without doubts, spending amount is very important to understand consumers’ consumption patterns and each consumer’s contribution to firms’ profits. Moreover, unlike the previous models, my data have information about the lifetime of each credit card. We know the exact lifetime of a credit card if it is cancelled before the end of sample period; or we know that the card is still valid at the end of the sample period. To analyze the data, I extend the model to accommodate these features.

With the data, I want to check the credit card consumption pattern both on aggregate and individual levels, and study the relationship between the consumption pattern and characteristics of the credit cards and their holders. By fitting the model to the data, I get the values of parameters which tell us the expected purchase frequency conditional on the activity of credit cards, expected spending amount of each purchase, and expected lifetime of each credit card. All of this information relates to the values of credit cards to the firm. To get a measure credit card’s contribution to the bank’s profits, I define the “value measure” of the credit card as the expected spending amount for the next month conditional on the
activity of the card, as the value measure of a credit card. To check whether the characteristics of credit cards and those of their users help identify the value of credit cards, I regress the parameters purchase frequency, dropout rate and the value measure on the characteristics of the credit card and those of its holders, including gender, education level, card level, promotion, and sale channel.

I have got some interesting results that are useful to improve customer relationship management. In aggregate, the expected lifetime of a typical credit card is estimated to be 101 months, expected number of purchases in a month for a typical credit card is 2.72, and expected spending amount in one purchase is 283 RMB (the currency in China, roughly equal to 45 dollars). On the individual level, I also get estimated values for the parameters, which reveal expected lifetime, expected number of purchases, and expected spending amount for each credit card. As a comprehensive value measure of credit cards, I also obtain the expected spending amount next month conditional that a credit card is currently active. The parameters and value measure for individual credit cards vary a lot, and they are good references to distinguish credit cards for marketing purpose. I also find that the characteristics of credit cards and their users show little information about credit card consumptions. To predict credit cards’ consumption in the future, past consumption record and the level of credit card are the most important information.

2. Literature Review

Based on the information about the recency and frequency of the purchases, several models have been developed to predict purchase patterns. Schmittlein et al. (1987) proposed the SMC model. The SMC model makes use of the recency and frequency of purchases by each customer, and is also referred to as the Pareto/NBD model. Key assumptions are the following: Purchases of each customer occur at random time and follow a Poisson process; the customers may drop out of the market randomly, and the lifetime of a customer follows an Exponential distribution; customers differentiate among each other with different parameters that guide the Poisson process and the Exponential distributions, and the parameters follow Gamma distributions.

Because the SMC model is difficult to implement, Fader et al. (2005) make an amendment to the model. Instead of assuming that the lifetime of a customer follows an Exponential distribution, they suppose that a customer drops out of the market after a purchase with probability \( p \), and stays in the market with probability \( 1 - p \) until the occurrence of the next purchase. To make the estimation of the parameters convenient, they assume a beta distribution for \( p \) to incorporate the difference across customers.

Abe (2009) extends the SMC model in two ways: First, he uses Markov chain Monte Carlo (MCMC) to estimate the model parameters. Second, he relates model parameters to customers’ characteristics information, such as income, education, and location, which help distinguish customers from each other.

There is also an amount of large literature on credit card consumption. Mansfield, Pinto and Robb (2012) conduct a meta-analysis on credit cards. As they say, “the purpose of this study is to review and integrate the literature surrounding consumer credit cards in various disciplines and offer a series of guiding research opportunities to the marketing discipline.” They review 537 scholarly articles and categorize them into the following areas: (1)
consumers’ emotional response to credit cards, (2) consumers’ credit card behavior which mainly concerns about numbers of cards owned, balance, access/availability, repayment, and credit card misuse, (3) consumers’ knowledge or beliefs about credit cards. However, they haven’t mentioned any studies on evaluating the values of credit cards from the perspective of credit card issuers. This study tries to fill the gap.

3. Extension of the Model to accommodate features in the data

3.1. Model

In this section, I will review SMC model’s assumptions that I will follow and the assumptions I have to incorporate spending amount into the model. Then, I will write out the likelihood function and the steps of simulations to get the parameters we want.

The assumptions are as follow. We adopts the assumptions on the SMC model that a purchase with a credit card occurs at a random time and follows a Poisson process and the credit card may be cancelled randomly and its lifetime follows an Exponential distribution. We propose that natural log of spending amount in each purchase follows a Normal distribution. Customers differentiate among each other with random parameters that guide the Exponential distributions, the Poisson distributions, and the Normal distributions. In more details, let’s first review Schmittlein’s assumptions

1. A credit card goes through two stages in its “lifetime”: it is active for some period of time, and then is cancelled and becomes permanently inactive.
2. While active, the number of purchases made by a credit card follows a Poisson process with transaction rate \( \lambda \); thus the probability of observing \( x \) transactions in the time interval \([0, t]\) is given by
   \[
P(X(t) = x|\lambda) = \frac{(\lambda t)^x e^{-\lambda t}}{x!}, \quad x = 0,1,2,3,\ldots
   \]
   And equally, the time between two purchases follows an Exponential distribution with \( \lambda \), and its probability density function is
   \[f(t_1 - t_{j-1}|\lambda) = \lambda e^{-\lambda (t_1 - t_{j-1}).}\]
3. Observed “lifetime” \( \tau \) of a credit card is Exponentially distributed with dropout rate \( \mu \). The probability density function is
   \[f(\tau|\mu) = e^{-\mu \tau}, \quad \mu > 0,\]
   and thus the probability that the lifetime is longer than \( T \) is
   \[R(T|\mu) = e^{-\mu T}.
   \]
4. Heterogeneity in transaction rates across credit cards follows a Gamma distribution with shape parameter \( r \) and scale parameter \( \alpha \) as
   \[g(\lambda|r,\alpha) = \frac{\alpha^r \lambda^r e^{-\lambda \alpha}}{\Gamma(r)}.
   \]
5. Heterogeneity in dropout rates across credit cards follows a Gamma distribution with shape parameter \( s \) and scale parameter \( \beta \)
   \[g(\alpha|s,\beta) = \frac{\beta^s \alpha^s e^{-\alpha \beta}}{\Gamma(s)}.
   \]
We further make the assumptions that

6. The natural log of the spending amount at \( j \)th purchase follows a Normal distribution as
\[ \log(y_j) \sim N(\mu_y, \sigma_y^2). \]

(7) Heterogeneity in the mean of the natural log of the spending amount, \( \mu_y \), across credit cards follows a Normal distribution as \( \mu_y \sim N\left( \mu_y^0, \left(\sigma_y^0\right)^2 \right) \), and diversity of inverse of the variance, \( 1/\sigma_y^2 \), across credit cards follows a Gamma distribution,

\[ 1/\sigma_y^2 \sim \left( \frac{1}{\sigma_y^2} \right)^{\delta_y+1} \exp \left( -\frac{\delta_y}{2} \frac{1}{\sigma_y^2} \right). \]

3.2. Model estimation

Likelihood:

We observe that purchases occur at time \( t_1, t_2, ..., t_x \), with corresponding spending amount \( y_1, y_2, ..., y_x \). If the card expires at time \( \tau \), the likelihood of observing the sample data given the parameters for the credit card is

\[
L(t_1, t_2, ..., t_x; \log(y_1), \log(y_2), ..., \log(y_x); \tau | \lambda, \mu, \mu_y, \sigma_y^2) = \lambda \exp(-\lambda t_1) \lambda \exp(-\lambda(t_2 - t_1)) \ldots \lambda \exp(-\lambda(t_x - t_{x-1})) \exp(-\lambda(\tau - t_x)) \times \exp(-\lambda(T - t_x)) \times \exp(-\mu T) \\
\times \frac{1}{\sqrt{2\pi} (\sigma_y)} \exp \left[ -\frac{1}{2} \frac{\left( \log(y_j) - \mu_y \right)^2}{\sigma_y^2} \right]
\]

Otherwise, if the card is still valid at time \( T \), the end of the sample period, the likelihood of observing the sample data given the parameters for the credit card is

\[
L(t_1, t_2, ..., t_x; \log(y_1), \log(y_2), ..., \log(y_x); T | \lambda, \mu, \mu_y, \sigma_y^2) = \lambda \exp(-\lambda t_1) \lambda \exp(-\lambda(t_2 - t_1)) \ldots \lambda \exp(-\lambda(t_x - t_{x-1})) \times \exp(-\lambda(\tau - t_x)) \times \exp(-\mu T) \\
\times \frac{1}{\sqrt{2\pi} (\sigma_y)} \exp \left[ -\frac{1}{2} \frac{\left( \log(y_j) - \mu_y \right)^2}{\sigma_y^2} \right]
\]

Simulate \( \lambda \) for each credit card:

The posterior distribution of parameter \( \lambda \), having observed recency and frequency of purchases, is

\[
L(\lambda | r, \alpha, x, \tau) \propto \lambda^r \exp(-\lambda r) \frac{\Gamma(r+1) e^{-\lambda x}}{\Gamma(r)} \propto \lambda^{x+r} \exp[-(\tau + \alpha) \lambda] \sim \text{Gamma}(x + r, \tau + \alpha),
\]

if the credit card is cancelled at \( \tau \). The posterior distribution of parameter \( \lambda \) is

\[
L(\lambda | r, \alpha, x, T) \propto \lambda^r \exp(-\lambda T) \frac{\Gamma(r+1) e^{-\lambda x}}{\Gamma(r)} \propto \lambda^{x+r} \exp[-(T + \alpha) \lambda] \sim \text{Gamma}(x + r, T + \alpha),
\]

if the credit card is valid at the end of sample period \( T \). We simulate \( \lambda \) from \( \text{Gamma}(x + r, \tau + \alpha) \) or \( \text{Gamma}(x + r, T + \alpha) \).

Simulate \( \mu \) for each credit card:
The posterior distribution of parameter \( \mu \), having observed the credit card is cancelled at \( \tau \), is
\[
L(\mu|s, \beta, \tau) \propto \mu \exp(-\mu \tau) \frac{\beta \mu^{e-1} e^{-\beta \mu}}{\Gamma(s)} \propto \mu^{s+1-1} \exp[-(\tau + \beta)\mu] \sim \text{Gamma}(s + 1, \tau + \beta).
\]
If the credit card is valid at the end of sample period \( T \), the posterior distribution is
\[
L(\mu|s, \beta, T) \propto \exp(-\mu T) \frac{\beta \mu^{e-1} e^{-\beta \mu}}{\Gamma(s)} \propto \mu^{s-1} \exp[-(T + \beta)\mu] \sim \text{Gamma}(s, T + \beta).
\]
We simulate \( \mu \) from \( \text{Gamma}(s + 1, \tau + \beta) \) or \( \text{Gamma}(s, T + \beta) \).

**Estimating parameters\( (r, \alpha) \):**

When we have simulated values of \( \{\lambda_i, i = 1, 2, ..., N\} \) for all credit cards, we estimate \( (r, \alpha) \). Because we assume that \( \lambda_i \) follows the Gamma distribution with parameters\( (r, \alpha) \), the joint likelihood of \( \{\lambda_i, i = 1, 2, ..., N\} \) is
\[
\prod_i g(\lambda_i|r, \alpha) = \prod_i \frac{\alpha^r \lambda_i^{r-1} e^{-\lambda_i} \alpha}{\Gamma(r)} = \frac{\alpha^{Nr} \prod \lambda_i^{r-1} \exp(-\alpha \Sigma \lambda_i)}{(\Gamma(r))^N}.
\]
We estimate \( (r, \alpha) \) by maximum likelihood.

**Estimating parameters\( (s, \beta) \):**

When we have simulated values \( \{\mu_i, i = 1, 2, ..., N\} \) of credit cards, we can estimate \( (s, \beta) \), the parameters in Gamma distribution which \( \mu_i \) follows. The joint likelihood of \( \{\mu_i, i = 1, 2, ..., N\} \) is
\[
\prod_i g(\mu_i|s, \beta) = \prod_i \frac{\beta^s \mu_i^{e-1} e^{-\mu_i \beta}}{\Gamma(s)} = \frac{\beta^{Ns} \prod \mu_i^{e-1} e^{-\beta \Sigma \mu_i}}{(\Gamma(s))^N}.
\]
We estimate \( (s, \beta) \) by maximum the likelihood.

**Posterior distribution of \( (\mu_y, \sigma_y^2) \) for each credit card:**

Because of diversity in the mean parameter, \( \mu_y \), and variance parameter, \( \sigma_y^2 \), of the natural log of spending amount among credit cards, we assume that for a randomly sampled credit card, \( \mu_y \) follows a Normal distribution and \( 1/\sigma_y^2 \) follows a Gamma distribution respectively. That is
\[
\mu_y \sim N \left( \mu_y^P, (\sigma_y^P)^2 \right),
\]
\[
1/\sigma_y^2 \sim \left( \frac{\delta_0}{2} \right) \exp \left( \frac{-\delta_1}{2} \right).
\]
The posterior distribution of mean \( \mu_y \) for a credit card, having observed its spending amounts, \( y_1, y_2, ..., y_x \) and given \( \sigma_y^2 \) known, is
\[
L(\mu_y|y_1, y_2, ..., y_x, \sigma_y^2) \propto \exp \left[ -\frac{1}{2} \frac{\Sigma \log y_i - \mu_y^2}{\sigma_y^2} \right] \times \exp \left[ -\frac{1}{2} \frac{(\mu_y - \mu_y^P)^2}{\sigma_y^P} \right]
\]
\[
\propto N \left( \frac{\Sigma \log y_i + \mu_y^P}{\sigma_y^P}, \frac{1}{\sigma_y^P + \sigma_y^P} \right).
\]
We simulate \( \mu_y \) for a credit card from \( N \left( \mu_y^P, \frac{\sigma_y^2}{\sigma_y^P + \sigma_y^2} \right) \).

The posterior distribution of the inverse variance\( 1/\sigma_y^2 \) for a credit card, having observed its spending amount, \( y_1, y_2, ..., y_x \) and given \( \mu_y \) known, is
\[
L(1/\sigma^2_y | y_1, y_2, ..., y_s, \mu_y) \propto \frac{1}{\sigma^2_y} \exp \left[ -\frac{1}{2} \frac{\sum (\log y_j - \mu_y)^2}{\sigma^2_y} \right] \times \left( \frac{1}{\sigma^2_y} \right)^{\frac{\delta_0}{2} - 1} \exp \left( -\frac{\delta_1 1}{2 \sigma^2_y} \right)
\]
\[
\propto \left( \frac{1}{\sigma^2_y} \right)^{\frac{\delta_0 + \delta_1 - 1}{2}} \exp \left( -\frac{\delta_1 + \Sigma (\log y_j - \mu_y)^2}{2} \right) \sim \text{Gamma} \left( \frac{\delta_0 + \delta_1}{2}, \frac{\delta_1 + \Sigma (\log y_j - \mu_y)^2}{2} \right).
\]

We simulate \(1/\sigma^2_y\) from \(\text{Gamma} \left( \frac{\delta_0 + \delta_1}{2}, \frac{\delta_1 + \Sigma (\log y_j - \mu_y)^2}{2} \right)\).

**Estimate parameters** \(\mu^p_y, (\sigma^p_y)^2\):

After we have simulated values of \(\mu_{y,1}, \mu_{y,2}, ..., \mu_{y,N}\), we estimate \(\mu^p_y, (\sigma^p_y)^2\) by maximizing the likelihood

\[
L(\mu^p_y, (\sigma^p_y)^2 | \mu_{y,1}, \mu_{y,2}, ..., \mu_{y,N}) \propto \frac{1}{(\sigma^2_y)^N} \exp \left[ -\frac{1}{2} \frac{\sum (\mu_{y,i} - \mu^p_y)^2}{(\sigma^2_y)^2} \right].
\]

**Estimate parameters** \((\delta_0, \delta_1)\):

After we have simulated values for \(\sigma^2_{y,1}, \sigma^2_{y,2}, ..., \sigma^2_{y,N}\), we estimate \((\delta_0, \delta_1)\) by maximizing the likelihood

\[
L(\delta_0, \delta_1 | \sigma^2_{y,1}, \sigma^2_{y,2}, ..., \sigma^2_{y,N}) = \prod_{i=1}^{N} \left( \frac{1}{\sigma^2_{y,i}} \right)^{\frac{\delta_0}{2} - 1} \exp \left( -\frac{\delta_1 1}{2 \sigma^2_{y,i}} \right).
\]

### 4. Empirical evidence

#### 4.1. Data summary

Table 1 is the summary statistics of credit card purchases. Our sample period is from 2008 to 2011, a total of 44 months, and the sample has consumption data for more than 10,000 credit cards. To make a preliminary analysis, we study the consumption data of the first 500 credit cards (data contains the consumption record of over 10,000 consumers and we use the first 500 consumers in the data for this analysis). Table 1 shows basic statistics of purchases with the credit cards. From the percentiles of the last purchase time, we can see that the median of the last purchase time is the 37th month and more than 25% credit cards have purchase records at the end of the sample period. From the statistics of the number of purchases, we find that the medium number of purchases is 57. Over 5% of the credit cards have more than 374 times of purchases. However, the number of purchases is less than 13 for more than 25% credit cards, which means many credit cards are only used occasionally. We also see that around 14% of credit cards are not used at all, and around 31% are cancelled during the sample period. For the credit cards that have purchase record, the medium of the natural log of spending amount of each transaction is 5.5, which means that, on average, one purchase costs around 245 RMB.

**Table 1. Simple statistics of credit cards**

<table>
<thead>
<tr>
<th>Percentile</th>
<th>5</th>
<th>25</th>
<th>50</th>
<th>75</th>
<th>95</th>
</tr>
</thead>
<tbody>
<tr>
<td>Most recent purchase time (in months)</td>
<td>0</td>
<td>13</td>
<td>37</td>
<td>44</td>
<td>44</td>
</tr>
<tr>
<td>number of purchases</td>
<td>0</td>
<td>13</td>
<td>57.5</td>
<td>162</td>
<td>374.5</td>
</tr>
<tr>
<td>ratio of number of cards having no purchase</td>
<td></td>
<td></td>
<td></td>
<td>0.142</td>
<td></td>
</tr>
</tbody>
</table>
ratio of number of cards canceled to total | 0.312  
lifetime for the canceled cards | 3  12  21.5  30  37.7  
average of the natural log of the spending amount for cards having purchase records | 4.3  5.1  5.5  6.1  7.1

4.2. Model estimation with the data

With the data, I proceed to estimate the model. First, initial values are given based on their relation to sample statistics of the data. For example, the expected time interval between two adjacent purchases is determined by $\hat{\lambda}_i$ as

$$E(t_{j,i} - t_{j-1,i}) = \frac{1}{\hat{\lambda}_i} \quad j = 1,2,\ldots, x_{i,t}, \text{ for credit card } i = 1,2,3,\ldots,N.$$  

Thus $\hat{\lambda}_i$ is given an initial value of

$$\hat{\lambda}_i = \frac{1}{\bar{y}(t_{i,j})} = \frac{1}{x_{i,t} \sum_{j=1}^{x_{i,t}} (t_{j,i}-t_{j-1,i})}.$$  

With the initial values of the parameters $\{(\lambda_i, \mu_i, \mu_{y,i}, \sigma_{y,i}^2), i = 1,2,\ldots,N\}$, we can simulate and estimate all parameters iteratively using the methodology illustrated in section 3.2. We observe the convergence visually. We do a total of 5000 iterations. Then, we discard the first 2500 iterations and use the second 2500 to get the estimated values of the parameters. When doing the estimations, we find that the estimated values of parameters $(\alpha, \beta)$ remain unstable even after many times of iterations. After careful check on the data, I find that for each credit card, we only have one observation about the lifetime, $\tau_i$ or $T_i$. In this situation, it is difficult to estimate both lifetime parameter $\mu_i$ and the parameters $(\alpha, \beta)$ that govern the heterogeneity in $\mu_i$ among credit cards. Thus, I simplify the model by giving $s$ a value of 1. In this case, the Gamma distribution reduces to an Exponential distribution. The estimated parameters are shown in Table 2 and Figures 1 and 2.

<table>
<thead>
<tr>
<th>Table 2. Estimated values of some parameters</th>
<th>$r$</th>
<th>$\alpha$</th>
<th>$\beta$</th>
<th>$\mu_y^\beta$</th>
<th>$(\sigma_y^\beta)^2$</th>
<th>$\delta_0$</th>
<th>$\delta_1$</th>
</tr>
</thead>
<tbody>
<tr>
<td>5 percentile</td>
<td>0.532</td>
<td>0.196</td>
<td>91.320</td>
<td>5.641</td>
<td>0.846</td>
<td>3.249</td>
<td>1.472</td>
</tr>
<tr>
<td>mean</td>
<td>0.558</td>
<td>0.205</td>
<td>101.249</td>
<td>5.647</td>
<td>0.852</td>
<td>3.505</td>
<td>1.634</td>
</tr>
<tr>
<td>95 percentile</td>
<td>0.583</td>
<td>0.214</td>
<td>111.638</td>
<td>5.653</td>
<td>0.859</td>
<td>3.768</td>
<td>1.802</td>
</tr>
</tbody>
</table>

Seeing the statistics in Table 2, we get a few important inferences at the aggregate level. We find that expected spending amount in one purchase for a typical customer is $\exp(\mu_y^\beta) = 283.44$ RMB from estimated value of $\hat{\beta}_y^\beta$, 5.647. From estimated value of $\beta$: $\hat{\beta} = 101.249$, we infer that the expected lifetime of a typical credit card is 101.249 months.

From estimated values of $r$, $\alpha$, we find that transaction rate of a typical credit card is $\frac{r}{\alpha} = 2.7219$, (because we have assumed the transaction rate follows a gamma distribution and mean of gamma distribution is $\frac{r}{\alpha}$) which means the expected number of purchases for one typical credit card in 44 months is $2.7219 \times 44 = 119$. We also see that the estimated values for parameters $\{(\lambda_i, \mu_i, \mu_{y,i}, \sigma_{y,i}^2), i = 1,2,\ldots,N\}$ in Figures 1, 2 vary a lot across
credit cards, and thus it is helpful to distinguish credit cards and manage “customer lifetime value” (CLV) on a systematic basis.

Figure 1. Estimated $\lambda$ and $\mu$ for each credit card (for the credit card that has no consumption record, we don’t estimate its $\lambda$, and give it a value of zero in the figure.)
Conditioned on credit card \( i \)'s survival into next month, the expected total spending amount in the month is
\[
\mathbb{E}(\text{number of purchases in one month} \times \text{spending amount each time}) = \\
\mathbb{E}(\text{number of purchases in one month}) \times \\
\mathbb{E}(\text{spending amount each time}) = \lambda_i \times \exp(\mu_{y,i} + \frac{\sigma_{y,i}^2}{2}).
\]

The probability that the holder does not cancel the credit card the next month given that it is active now is
\[
\exp(-\mu_i).
\]

Thus the natural log of the expected spending amount with credit card \( i \) the next month conditional that it is active now is
\[
\log Q_i = \log \left( \exp \left( \mu_{y,i} + \frac{\sigma_{y,i}^2}{2} \right) \times \lambda_i \times \exp(-\mu_i) \right) = \mu_{y,i} + \frac{\sigma_{y,i}^2}{2} + \log(\lambda_i) - \mu_i.
\]

Having estimated the parameters in the model, I calculate natural log of the expected spending amount \( \log Q_i \), and use it to measure the value of each credit card. The higher the measure, the higher value the credit card has. I calculate the value measure and plot it in Figure 3. For the credit cards that have no consumption record, I don’t estimate some of the parameters and assign their value measures as 0. The value measures take quite different values for different cards, and can help us distinguish credit cards for marketing.
4.3. Test of the model assumptions

In our model, many assumptions are taken from previous studies and are taken as granted. For example, we follow the SMC model’s traditional assumptions that purchasing times of each credit card follow a Poisson distribution and the heterogeneity in purchase frequency across customers follow a Gamma distribution. We also propose new assumptions that the natural log of spending amount of each purchase follows a Normal distribution for each credit card and heterogeneities in means and variances of the Normal distributions across credit cards are described respectively by a Normal distribution and an inverse Gamma distribution. From Bayesian statistics, the distributions describing heterogeneities in means and variances come from the conjugate priors to make simulation simple. To demonstrate that the assumptions are reasonable and fit the data well, I test whether the spending amounts of all customers follow a Normal distribution here.

I perform a chi-square goodness-of-fit test for each credit card. The null hypothesis is that the spending amount of each purchase in the data follows a normal distribution at the 5% significance level. The results are reported in Figure 4. The null hypothesis is not rejected when H0 takes a value of 0; otherwise, H0 takes a value of 1 meaning that the normal distribution is rejected. P value is the probability of observing the given result if the null hypothesis is true. We see that fewer than one half credit cards are shown to follow Normal distributions. After a careful check, I find a possible reason for such a large number of rejections. Our data is observed monthly. We use monthly spending amount divided by the number of purchases in the month to represent the spending amount of each purchase. Thus, a consumer’s spending amount of each purchase in the same month is uniformly distributed. Even if the actual spending amount follows a normal distribution, the data behaves less likely to be from a normal distribution. Finally, although the real distribution of purchase amount for a credit card may be complicated, normal distribution is still a good approximation if we focus on first and second moments of the distributions.
5. Relation between credit card purchase and characteristics of credit card and its holder

From the consumption data, I have estimated the model parameters and calculated the value measures of credit cards. I now check on the relation between them and characteristics of credit cards and their holders.

Table 3 shows the relation between purchase frequency parameter $\lambda$ and characteristics of credit cards and their holders. Gender has no effect on purchase frequency. Education level\(^1\) has some effect, and highly educated credit card holders tend to buy less frequently. Card level\(^2\) has a significant effect on purchase frequency. Holders of a level-2 card purchase more frequently. Level-2 cards are “gold cards.” Only consumers with very good credit scores can apply for a gold card successfully.\(^3\)

---

\(^1\) Education has 5 levels. Edu\(_h\) means that the holder graduated from high school. Edu\(_p\) means the holder graduated from college, and edu\(_u\) means that holder graduated from university. Edu\(_m\) means the holder has a master degree or above. The base level is holder who graduated from middle school or below.

\(^2\) Level-1 card is called common card. Level-2 is gold credit, which is the most premium one among the three. Level-3 card is called youth card for young people.

\(^3\) Baseline education means the card holder graduated from middle school. Edu\(_h\) means the card holder graduated from high school. Edu\(_p\) means the holder graduated from junior college. Edu\(_u\) means the holder graduated from college. Edu\(_m\) means the holder has a master degree or PhD degree.

\(^4\) Baseline promotion is the promotional event for common and gold card in 2007. Promotion\(_2\) is the promotional event for youth card in 2007. Promotion\(_3\) is the promotional event for common and gold card in 2008. Promotion\(_4\) means the card was open without any promotions.
Definition of variables:

Education has 5 levels. Edu_h means that the holder graduated from high school. Edu_p means the holder graduated from technical schools, and edu_u means that holder graduated from university. Edu_m means the holder has a master degree or above. The base level is holders who have education level of 9th grade or below.

Level-1 card is called common card. Level-2 is gold cold, which is the most premium one among the three. (Banks usually only give gold card to customers with high incomes and good credit scores.) Level-3 card is called youth card for young people (people who are over eighteen but has no checking or saving accounts can apply for youth cards under the accounts of legal parents).

Baseline promotion is the promotional event for common and gold card in 2007. Promotion_2 is the promotional event for youth card in 2007. Promotion_3 is the promotional event for common and gold card in 2008. Promotion_4 means the card was open without any promotions.

Gender is a dummy variable. Male is assigned to be 1 and female is assigned to be 0.

Age variable is the age at which the consumer opened the card.

Baseline channel level means the cards were opened at the branch banks. Channel_1 means the cards were opened through third-party stores. Channel_2 means the card were opened directly at the bank.

Table 3. Regressing $\lambda$ on characteristics of credit card and its holders

<table>
<thead>
<tr>
<th></th>
<th>constant</th>
<th>2.63</th>
<th>3.98</th>
<th>2.21</th>
<th>2.82</th>
<th>1.63</th>
<th>3.58</th>
<th>2.59</th>
<th>3.24</th>
</tr>
</thead>
<tbody>
<tr>
<td>gender</td>
<td>0.16</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.50)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>edu_h</td>
<td>-2.80</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(-1.51)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>edu_p</td>
<td>-1.44</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(-1.98)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>edu_u</td>
<td>-1.84</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(-2.63)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>edu_m</td>
<td>-0.81</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(-1.18)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>card_level2</td>
<td>1.75</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(5.00)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>card_level3</td>
<td>-0.06</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(-0.13)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>promotion_2</td>
<td>-0.66</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(-1.57)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>promotion_3</td>
<td>1.66</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(-0.47)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>promotion_4</td>
<td>1.74</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(1.21)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>channel_2</td>
<td>-1.01</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
\( \mu_y \) is the expected natural log of spending amount for each purchase. Table 4 reports the evidence of regressing \( \mu_y \) on the characteristics of credit cards and their holders. We see that gender, education level, and promotion have no obvious effect on the parameter. Card level has a positive effect, and sale channel has a minor effect. Gold cards (level 2) and youth cards (level 3) have higher spending amount for one purchase on average than level-1 cards, which are common cards.

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>( \mu_y )</th>
<th>( \text{R}^2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td>0.24</td>
<td>0.00</td>
</tr>
<tr>
<td>Edu_s</td>
<td>-0.62</td>
<td>0.90</td>
</tr>
<tr>
<td>Edu_h</td>
<td>-0.74</td>
<td>0.81</td>
</tr>
<tr>
<td>Edu_p</td>
<td>-0.04</td>
<td>1.38</td>
</tr>
<tr>
<td>Edu_u</td>
<td>1.34</td>
<td>0.80</td>
</tr>
<tr>
<td>Card level 2</td>
<td>0.32</td>
<td>1.00</td>
</tr>
<tr>
<td>Card level 3</td>
<td>0.81</td>
<td>0.94</td>
</tr>
<tr>
<td>Promotion 2</td>
<td>-0.91</td>
<td>0.46</td>
</tr>
<tr>
<td>Promotion 3</td>
<td>-0.64</td>
<td>0.51</td>
</tr>
<tr>
<td>Promotion 4</td>
<td>-0.01</td>
<td>0.08</td>
</tr>
<tr>
<td>Age</td>
<td>0.01</td>
<td>0.03</td>
</tr>
<tr>
<td>R-squares</td>
<td>0.00</td>
<td>0.08</td>
</tr>
</tbody>
</table>
Finally I regress the credit card value measure $\log Q$, on the characteristics of credit cards and their holders and find the relation between them in Table 5. The evidence shows that card level affects the value measure. On average, gold cards have the highest value and youth cards contribute more value than common cards.

<table>
<thead>
<tr>
<th>Table 5. Regressing $\log Q$ on characteristics of credit card and its holders</th>
</tr>
</thead>
<tbody>
<tr>
<td>constant</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Gender</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>edu_s</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>edu_h</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>edu_p</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>edu_u</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>card_level2</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>card_level3</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>promotion_2</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>promotion_3</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>promotion_4</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>channel_2</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>channel_3</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>age</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>R-sqaures</td>
</tr>
</tbody>
</table>

6. Conclusion

In the data on credit card consumption, number of purchases and total spending amount each month is observed. To make use of observations of spending amount, a critical information to evaluate credit card consumptions, I extend the models proposed by Schmittlein et al. (1987) and Abe (2009). I assume that natural log of spending amount of each credit card follows a Normal distribution, and diversity in the means and variances of the normal distributions across credit cards follow a Normal distribution and an inverse
Gamma distribution respectively. With the data, I get reliable estimated values of the parameters, which provide important information about the consumption pattern of each credit card and credit cards on aggregate level. For example, in aggregate, the expected lifetime of a typical credit card is estimated to be 101 months, expected number of purchases in a month is 2.72, and expected natural log of spending amount in one purchase is 6.4. The expected frequency and expected spending amount of each purchase varies a lot across credit cards. To evaluate credit cards with these estimated parameters, I introduce the expected spending amount for the next month conditional that the credit card is active now as a value measure of the credit card. The value measure varies much across credit cards and thus is helpful to distinguish values of credit cards. Finally by regressing the estimated parameters and the value measure on the characteristics of credit cards and their holders, I find that card level has a significant effect on credit card purchase patterns and other characteristics have no significant effects.

References
Abe Makoto. 2009. “Counting your customers” one by one: a hierarchical Bayes extension to the Pareto/NBD Model, Marketing Science, 28(3), 541-553.