Consumer Credit Risk and Spending Behavior

Ali Bakhtiari
B.P.S. Murthi

The University of Texas at Dallas
Jindal School of Management

Abstract

Since about a decade ago, most banks started closely watching consumer spending behavior to identify those who are buying from categories that can signal financial distress situations. In case of continued spending on certain categories, banks would adjust APR or credit limit to reduce the future loss due to possible delinquency. Despite the extensive body of research on consumer credit risk prediction and its implications for bank's risk management decisions, there is no literature considering the relationship between consumer spending behavior and credit risk. In this research, we propose a hidden Markov model (HMM) to dynamically predict consumers' credit risk and classify them based on their transitioning between different risk levels. Using historical repayment behavior and current-period credit spending amount at different categories as predictors and account delinquency status as response variable, we predict the transition probability between three hidden levels of high, low, and no risk. We find that buying from discount stores and pawnshops, spending excessively on entertainment such as gambling, dating services, bowling alleys and video games, and paying off utility bills or governmental fines are significantly associated with higher credit risk. However, using credit card for necessity items such as grocery shopping, gas and service stations are linked to lower levels of risk. We also use cluster analysis to group consumers based on their risk evolution patterns. We find four risk evolution patterns: always-high, always-low, moderate-increasing, moderate-decreasing. We show that consumers with moderate-decreasing risk level are the most profitable customers for the bank as they pay higher finance charges than always-low and increasing-moderate groups but also have lower probability of delinquency than always-high consumers. We also show that always-high consumers use credit card mostly for paying bills, utilities, government fines and fees, whereas always-low consumers use credit card for pricier purchases such as home furniture, durables goods, and car. Consumers with moderate-increasing risk pattern are the most frequent users of credit card as they have highest share of spending in most of the purchase categories. Our paper contributes to the existing literature in mainly two ways. We use HMM for credit risk prediction which also yields dynamic classification of consumers by their credit risk much simpler than other methods. We are also the first paper to consider the relationship between consumer credit risk and purchase categories.

1 Introduction

In the consumer credit card industry, the profit of credit issuers depends heavily on the balance repayment behavior of customers which, in turn, relates to their credit risk. The amount of risk is related to whether a customer pays off the balance regularly, misses payments, or defaults. The ability of credit providers to distinguish low-risk delinquent customers from high-risk ones has significant impact on their profitability (Zhao et al. 2009). As a common practice for decades, most of the credit providers relied on credit scoring systems such as FICO to rank customers based on their credit worthiness. Most of these systems use cardholders' credit bureau data containing their past credit history. For instance, FICO score includes factors such as utilization rate, payment history, and length of credit history. Despite the ubiquitous application of credit scoring methods by the credit card providers, they are not capable of capturing the dynamics of existing customers.

Since consumer behavior is evolving through time, behavioral scoring is an obvious extension to the existing credit scoring methods (Thomas 2005). During the past couple of decades, some of the major banks started adopting behavioral scoring methods and used them alongside the credit scoring information (Li and Zhong, 2012). In behavioral scoring, banks use repayment and purchase behavior of customers to provide a more accurate evaluation of default risk. Behavioral scoring methods can be categorized into two general approaches (Thomas, 2000). One approach uses the same static methods as credit scoring but with some additional (and mostly aggregate) variables derived from customer repayment behavior. The other approach uses some form of statistical- or probability-based models to dynamically predict default risk based on the customers' most recent repayment behavior. Most of the advancement in credit risk prediction are resulted from the latter approach as it directly takes into account the dynamic nature of consumer behavior.

Our approach to model the dynamic credit card behavior of customers is structurally different from the existing approaches. The difficulty with capturing consumer behavior dynamics is that the nature and structure of dynamics is often latent. One could model such dynamics, by allowing consumers to transition over time between a set of discrete states (Netzer et. al 2009). To overcome the problem of unobserved states one could describe a set of latent states and transitions between these states and translate these latent states to the observed behavior through a stochastic model. Following this approach, we propose a hidden Markov model (HMM) to estimate the transition probabilities among latent risk states and their effect on probability of default. In our model, the transition between risk states is a function of time-varying covariates such as repayment behavior and purchase categories. These covariates can, in turn, affect the delinquency status by shifting consumer to a different (unobservable) risk state. MacDonald and Zucchini (1997, Chapter 4) describe several applications of HMMs in areas ranging from biology, geology, and climatology to finance and criminology. In econometrics, Hamilton (1989) proposed an HMM to estimate the impact of discrete regime shifts on the growth rates of the real gross national product. Within the marketing literature, HMMs are closely related to the family of latent class models (Kamakura and Russell 1989). Like most latent class models, HMMs classify individuals into a set of states or segments based on their credit card behavior. However, unlike the latent class models, in HMMs the membership in the latent states is dynamic and follows a Markov process.

Our proposed HMM contributes to the literature of consumer credit risk assessment from several aspects. First, using HMM we can capture the dynamics of repayment and purchase behavior with less restrictive assumptions than those of proposed state space models (Zhao et. al 2009). Second, separation of low- and high-risk customers is a natural result of HMM whereas in survival analysis inferring the risk type is not straightforward. Moreover, HMM also encompasses one of the main reasons of using survival analysis for credit risk evaluation. Similar to survival analysis, censoring can be accommodated into HMM as individuals only contribute to the risk prediction for as long as they are observed and do not contribute to calculations of transition from state they are in when they are censored. In fact, using a Markov model for right-censored data has an advantage over survival analysis in that each censored subject contributes more information to the model than it can contribute to survival analysis with one end point that the subject did not reach. It is because the prior state transitions of individuals add useful information (Hillis, et. al 1986). Third, it is important to understand the drivers of the dynamics rather than merely build a model that fits the dynamics in the data. While many of research in this area is focused on just predicting default without explaining what is contributing to it, we use time-varying covariates in estimation of transition probabilities. Fourth, we take into account the information regarding consumer spending categories and relate them to the risk of default.

To our knowledge, repayment behavior-based variables are widely used in research studies on behavioral scoring and only a handful of papers investigated the implications of purchase category on one's future credit risk. Since about a decade ago, some of the major credit card issuers started creating credit card profiling systems using consumers' amount, time, location, and category of credit transactions to form spending habit categories[1]. The idea is that credit spending behavior can signal financial distress situation and tell a lot about consumer's ability to repay the outstanding balance. For instance, using cash advance, buying from discount stores or pawnshops, spending excessively on casinos, dating services and video games can be early warnings of elevated risk. According to a congress report published in 2010[2], during the period of 2006 to 2009 some of card issuers used broad category definitions to apply risk management measures which affected changes of credit terms for about 340K

[1] http://www.creditcards.com/credit-card-news/credit-card-profiling-fed-report-1282.php

[2] http://www.federalreserve.org, *Report to the Congress on Reductions of Consumer Credit Limits Based on Certain Information as to Experience or Transactions of the Consumer*

customers. It should be noted that such credit risk profiling practices are mainly made possible by the increased data storage capacity and processing power of computers. Despite the prevalence of credit card profiling, very few research in the consumer credit risk literature has investigated the existence of such relationship between purchasing habits and credit risk. In this study, we are looking at one of the aspects of purchasing behavior and leave the rest for future research. Specifically, we examine if consumer credit risk has any association with the purchase categories to see if customers at different degrees of credit risk tend to purchase more from certain categories. Vissing-Jorgensen (2011) studies the determinants of default risk on a given loan based on customer shopping behavior. She found that differences in default losses are significantly different across product categories after controlling for standard measures of risk. She argues that product mix purchased should be an important component of credit scoring. This study is focused on loans which implies that it has considerable differences from the credit cards in terms of risk evaluation methods. The risk of default is less dynamic for loans compared to credit cards because loans have a certain maturity date and installments are mostly predetermined whereas in credit cards debt can be constantly changing due to consumer behavior dynamics. Despite the differences of risk between loans and credit cards, we also find that customers belonging to different credit risk categories tend to buy from different group of products. We also evaluate the effect of certain purchase category on transitioning customer between different risk states.

In this paper, we use hidden Markov model to estimate risk level at each point of time and use the estimated risk levels to classify customers based on their evolution across different risk levels. Using our dynamic model, we identify four groups of customers based on their risk evolution pattern and show that customers at different risk groups purchase from different categories. Our paper contributes to the current literature in multiple ways. We are the first paper to apply hidden Markov model to dynamically predict risk and classify customers based on their transitioning pattern across different risk levels, and show the relationship between purchase category and credit risk. We find that the impact of different repayment covariates (e.g. utilization rate) on credit risk are the same, however, the magnitude of impact differs based on the current hidden risk state. We identified four clusters of customers based on their risk state evolution through time. Interestingly, each cluster shows different characteristics in terms of risk evolution and potential profitability for the bank. We find that the cluster of customers who start with high utilization rate and high risk states but evolve into lower risk states are the most profitable for the bank because they pay the highest finance charges while having a modest, decreasing level of risk. Also, we find that there is a segment of customers who starts with high utilization and delinquency rate and stay in the same state. They carry large balances and have a fairly high likelihood of default and should be the first priority of bank's risk management measures. We also show that the highest spending categories are totally different among different risk clusters.

The remainder of this paper is organized as follows. Section 2 reviews the literature on consumer credit risk prediction methods. In section 3, we develop our hidden Markov model and describe the estimation process. Section 4 describes the data and reports the results of estimation of our model on the credit card transaction data. We also discuss how different risk clusters are associated with different category purchases. In Section 5 we conclude the paper and suggest future research directions based on current work.

2 Literature Review

2.1. Credit Risk Assessment

Research studies on consumer credit risk assessment mainly fall into one of the following categories: (1) credit worthiness scoring for identifying high-risk credit card applicants from low-risk ones, (2) behavioral scoring for managing the risk of current customers by assessing their future risk based on their past credit behavior and evaluating the profitability of customers mainly based on their default likelihood (see Thomas 2000; Thomas et al. 2005; Thomas et al. 2001 for a review on research in this category). The proposed method in this research belongs to the latter category. The most common scoring approach used by both researchers and credit providers is Logistic Regression (LR) (Stepanova and Thomas, 2002). Using this method, one would specify a period of time (e.g. 24 months) and fit a LR model to the historical customer data for predicting the probability of default as a function of the applicant's credit bureau information. Lower predicted probability of delinquency means better credit worthiness for that customer. Banks usually set a cut-off threshold and approve higher credit and lower annual percentage rates (APRs) for those customers whose predicted probability is less than the cut-off point.

Survival analysis is an alternative to LR that is reasonably simple, in that it does not involve an overly parameterized model. In survival analysis for scoring, the objective is to model the distribution of the time to default (or of some other event associated with default), as opposed to the LR objective of predicting the probability of default within a single, pre-specified period of time. In survival model, the distribution of response variable is allowed to be a function of applicant's credit variables via a proportional hazards (PH) function (see Leemis et. al (2000) for comprehensive treatments of PH and other survival models). For the sole purpose of predicting the probability of default within a single specified period, traditional PH survival modelling has little advantage over LR. Stepanova and Thomas (2002) have found their performances to be nearly indistinguishable.

Different variations of survival models (Crook 2009, Stepanova and Thomas 2002, Glennon and Nigro 2005, Ricardo et. al 2010, Zhang and Thomas 2012) are widely used by both researchers and credit providers. The survival analysis concentrates on the time dependency of the default risk alone, not on the delinquency states leading to it, therefore, separation of low-risk from high-risk customers is not straightforward. Amongst other methods of behavioral scoring, are the AI-based algorithms (Examples of such methods are support vector machines by Khandani et al. 2010, neural networks by Pacelli and Azzollini, 2011, and Bayesian network classifiers by Pavlenko and Chernyak, 2010) have also been developed. These methods attempt to fit more complex models with higher degrees of nonlinearity between the predictors and credit risk. One of the objections (Capon 1982) to some of the credit risk prediction methods[3] is that they are all focused on which customers are going to default without providing any explanation on why they default.

2.2. Behavioral Credit Risk Influencers

By accounting for the inherent dynamics exist in consumer credit risk, behavioral scoring methods make it possible for the banks to apply risk management policies more effectively. In most of the papers that used behavioral-based scoring models (see Thomas 2000, Hsieh 2004, Lin & Zhong 2012) to predict credit risk, repayment behavior of customer is used alongside the credit scoring factors such as demographic characteristics. Using customer repayment behavior information, researchers are trying to find patterns in consumers' credit card behavior and group those with similar behavioral pattern. It is crucial for the banks to know which segments of their customers are low risk and are going to pay back their balance in full, and which segments are low risk and going to revolve balances and generate revenue through interest payments. Similarly, it is important for them to identify which segments are high risk and likely to default. According to credit card industry reports (Zhao et. al 2009), 40% of cardholder pay their balances in full and only contribute to the bank revenue through interchange fees. High risk customers account for about 4% (varies depending on economic status) and provide almost no revenue for the bank and in most cases they are not able to pay

[3] Mostly AI-based methods such as decision tree or artificial neural network

back and banks write off their balance as bad debt. Revolvers are the largest and most profitable segment accounting for almost 60% of cardholders and 70% of banks' revenue. By using behavioral scoring and classifying customers into different groups, banks are able to classify customers by their level of risk and apply their risk/profit management policies more effectively. Behavioral scoring is especially a useful method because it takes into account the learning process in repayment behavior and customers' evolution in better managing their credit score status.

Repayment behavior can be captured by metrics showing the amount of repayment and number of times a payment is missed. This kind of data is rarely available to the researchers which may explain the limited academic research in this area which have used any measures of repayment behavior. Amongst the few who had done so is the research by Zhao et. al (2009) which use repayment behavior metrics to predict customer profitability. They propose a Type II Tobit model to predict the risk type of customers. Their model accounts for customer oversight and distinguishes between low- and high-risk delinquent customers. They argue that customers with higher estimated repayment ratio (payment over total balance) have lower risk and their delinquency are more likely to be due to oversight rather than the inability or unwillingness of customer to repay the debt. Although, their state space model accounts for both consumer dynamics and heterogeneity, the cut off value to classify the high- and low-risk customers needs to be imposed by the researcher and is not a consequence of the modeling approach. Another limitation of the observed states models is their restrictive account for consumer behavior dynamics whereby an ad-hoc specification of state dependence is added to an otherwise static model.

There are also a limited number of papers that account for current-month credit spending as a predictor of credit card default. However, to the best of our knowledge, there is no paper which has looked at the relationship between purchase categories and credit card default probability. Information on purchase category is collected by the banks whenever a customer makes a transaction with the credit card. A category is assigned to each transaction based on the Merchant Category Codes (MCC) to keep track of different types of consumer spending. Depending on the credit card issuer, MCC is a list of around 600-700 different categories which makes it very difficult to be used as predictors in a statistical model due to the large number of parameters that needs to estimated. Moreover, possible correlation between categories also adds to the number of parameters. In addition to that, a broader purchase category definition would do better in terms of interpretability of the results. In our research, we use MCC categories from both types of discretionary and non-discretionary spending as predictors in our model use the results to see which of the categories are consumed by the low and high risk group of customers. According to Danziger (2004), discretionary spending categories fall into four basic categories: utilitarian purchases, indulgences, lifestyle luxuries, and aspirational luxuries. In utilitarian purchases customers need the item because they perceive that it makes their lives better in a meaningful way. Most of the home appliances such as blender, microwave oven, food processors, and water purifiers (to name a few) fall into this category. Indulgences are life's little luxuries that consumers can buy without guilt. Examples of such purchases include cosmetics, flowers, perfumes, entertainment, video games, computers and electronics. Lifestyle luxuries are the purchases that have a practical aspect as well. Buying a car, or a watch are examples of that. Finally, aspirational luxuries include purchases that owning them bring pure joy such as buying an original art or piece of antique, and branded products. Non-discretionary categories, on the other hand, are the necessary spending that people cannot live without or have to be purchased. Examples include food (made in home), gas, health, bills and payments, transportation fees, and educational expenses. For our research, we choose few categories from both discretionary and non-discretionary spending to keep the model parsimonious. Specifically, from non-discretionary MCC, we group grocery shopping, gas, service station, dairy product into Necessity Goods category. In terms of risk-related categories from discretionary spending we follow the suggestions from popular articles and congress report on credit card profiling[4]. The congress report shows that entertainment activities when done excessively can increase the risk of default. We combine spending on MCCs for gambling, casinos, billiard pool, bowling alleys, dating and escorting services, and video games into Entertainment category. Also, buying from pawnshops, used merchandise stores, one-dollar stores are a signal of financial distress[5]. We combine them under Discount Store category. In addition to repayment behavior, demographic characteristics of customer are also used in behavioral scoring research. Using a data from a Malaysian bank, Teoh et. al (2013) show that age, income, and marital status have significant correlation with credit card spending behavior whereas occupation is not significantly related to it. Kao et. al (2013) proposed a Bayesian behavioral scoring model to identify the factors that affect default risk. They find that high-income, female, or cardholders with higher education are more likely to have good repayment ability. Gender and income information is not available in our data and we use age, occupation, and card type (reward card or not) as the demographic variables in our model.

3 Model

In this section we develop a hidden Markov model (HMM) that predicts the credit risk level of a customer using information on past repayment behavior and purchase categories across several time periods. The response variable of the model is debt repayment status defined as a categorical variable which takes the value of 1 if customer misses two (or more) consecutive payments (hereafter referred to as delinquent status) and 0 otherwise. We are not considering

---

[4] http://www.federalreserve.org, *Report to the Congress on Reductions of Consumer Credit Limits Based on Certain Information as to Experience or Transactions of the Consumer*

[5] http://www.nytimes.com/2009/05/17/magazine/17credit-t.html?_r=0, *What Does Your Credit Company Know About You?*

one-time payment misses for a couple of reasons. First, the focus of our model is to capture occasions where missing a payment is very likely going to turn into charge-off and bad debt for the banks. Credit issuers usually find their most profitable segments amongst the customers who revolve balance and may occasionally miss a payment. Second, missing one payment can also be attributed to customer oversight which doesn't necessarily have anything to do with an elevation in credit risk. A customer's risk level increases only if multiple payments are missed in a certain periods of time. Therefore, we consider balance repayment over longer periods of time than one. In the next section, we set up the model and explain how we relate the repayment behavior and spending categories to account debt repayment status.

3.1. *Model Setup*

We denote customer i's, $i \in \{1, \dots, N\}$, debt repayment status at time t, $t \in \{1, \dots, T\}$, by $\boldsymbol{Y}_i$ which is a vector of categorical response variable takes the values $\{0,1\}$ representing delinquent/non-delinquent status. Also let $\boldsymbol{Y}$ be the long-format vector of debt repayment status of all individuals with the size of $NT \times 1$ and made by stacking up N individual vectors, $\boldsymbol{Y}_i$. We include time-varying covariates in addition to individual-specific variables to account for the impact of dynamic factors on account delinquency. In our model setup, the probability distribution of response variable, debt repayment status, is conditional on the hidden states which, in turn, are dependent on the covariates. We consider three types of covariates including historical (previous period) repayment behavior variables, the amount of purchase in current period in different categories, and customer demographics. Repayment behavior covariates are denoted by $\boldsymbol{X}$ which is a matrix of size $NT \times p$ capturing repayment behavior of customers in the last time period, $\tau - 1$, where p is the number of covariates and τ is the current time period. Following the Markov process assumption of our model, we don't consider time periods before $\tau - 1$. The amount of purchase in different categories at current time period, τ, is denoted by

$\boldsymbol{W}$ which also is matrix of size $NT{\times}q$, where q is the number of categories. Finally, customer demographics are denoted by $\boldsymbol{Z}$ which is matrix of size $N{\times}r$ where r is the number of demographic covariates.

The general formulation of hidden Markov model assumes the existence of hidden states which are related to each other through a first-order Markov process and a response variable which its probability distribution is conditionally dependent on the hidden states. This assumption implies that response variables are conditionally independent given the hidden states and, therefore, there is no serial correlation between the vectors of response variable (Rabiner 1989). Netzer et. al (2009) used similar formulation to model the customer relationship dynamics. This assumption simplifies the model by letting the current state be only dependent on previous state, which itself is a natural outcome of Markov process assumption within the hidden states. The conditional distribution of debt repayment status at any time period can be written as

$$y|s = p(\boldsymbol{Y}_i^{(t)} = y_{it}|\boldsymbol{S}_i^{(t)} = s_{it}) \tag{1}$$

Where $\boldsymbol{S}_i$ is the vector of hidden risk states for customer i, and y and $s, s \in \{1, \dots, k\}$, are the realization of response and hidden states vectors at time t, and k is number of hidden states. At each time period customers are transitioning (or staying) to different hidden risk states with a probability where the probability of transitioning between states depends on the vector of covariates, $[\boldsymbol{X}, \boldsymbol{W}, \boldsymbol{Z}]$. We use multinomial logit framework to calculate the transition probabilities between different states. Hidden risk states are the response variables of the logit model, therefore, the log-odds of transitioning from one state to another can be written as,

$$\pi_{s\bar{s}} = log\ \frac{p(\boldsymbol{S}_i^{(t)} = s|\boldsymbol{S}_i^{(t-1)} = \bar{s}, \widetilde{\boldsymbol{Q}}_i^{(t)} = q_{it})}{p(\boldsymbol{S}_i^{(t)} = \bar{s}|\boldsymbol{S}_i^{(t-1)} = \bar{s}, \widetilde{\boldsymbol{Q}}_i^{(t)} = q_{it})} = \alpha_{0s} + \boldsymbol{x}'\boldsymbol{\beta}_s + \boldsymbol{w}'\boldsymbol{\gamma}_s + \boldsymbol{z}'\boldsymbol{\lambda}_s \qquad \begin{matrix} t \geq 2 \\ s \neq \bar{s} \end{matrix} \tag{2}$$

Where $\pi_{s\bar{s}}$ is the log-odds of transitioning from one state to another, $(\boldsymbol{\beta}_s, \boldsymbol{\gamma}_s, \boldsymbol{\lambda}_s)$ are covariate parameters to be estimated and $\widetilde{\boldsymbol{Q}}$ denotes the joined matrix of covariates, $[\boldsymbol{X}, \boldsymbol{W}, \boldsymbol{Z}]$. For state probabilities at $t = 1$, we draw randomly from a Dirichlet distribution. We re-parameterize the model in (2) as the difference between two sets of parameters for a more parsimonious model,

$$\pi_{s\bar{s}} = \alpha_{0s} + \boldsymbol{x}'(\boldsymbol{\beta}_{\bar{s}} - \boldsymbol{\beta}_s) + \boldsymbol{w}'(\boldsymbol{\gamma}_{\bar{s}} - \boldsymbol{\gamma}_s) + \boldsymbol{z}'(\boldsymbol{\lambda}_{\bar{s}} - \boldsymbol{\lambda}_s) \tag{3}$$

Where $\boldsymbol{\beta}_1 = \boldsymbol{\gamma}_1 = \boldsymbol{\lambda}_1 = \boldsymbol{0}$ for identification purposes. Having specified the conditional response and hidden states transition probabilities, we can write the joint probability distribution of response variable as,

$$p(y_{i1}, \ldots, y_{iT}) = \pi_{1s} \prod_{t=2}^{T} \pi_{s\bar{s}} \prod_{t=1}^{T} p(y_{it} | \boldsymbol{S}_i^{(t)} = s) \tag{4}$$

3.2. *Estimation Method*

Given a sample size of N, the log-likelihood function can be written as follows,

$$LL(\boldsymbol{\theta}) = \sum_{i=1}^{N} \log\, p(\boldsymbol{Y}_i | \widetilde{\boldsymbol{Q}}_i) \tag{5}$$

Where $\boldsymbol{\theta}$ is the vector of parameters, $\boldsymbol{\theta} = [\boldsymbol{\beta}_s, \boldsymbol{\gamma}_s, \boldsymbol{\lambda}_s]$. To estimate the parameters, we use Expectation-Maximization method which is a computationally-efficient algorithm suggested by Baum et al. (1970) and Dempster et al. (1977) to find the maximum of likelihood function, especially when the model depends on the latent variables. Expectation Maximization algorithm iterates between two steps of Expectation (E-step) and Maximization (M-step) to find the parameters that maximize the log-likelihood function. It starts by an initial guess for the parameter values and then in the E-step the expected value of the log-likelihood function is calculated by integrating on the conditional distribution of the observed and latent variables given the current parameter estimates. Next, in the M-step, the parameter values maximizing the expected log-likelihood are estimated. These parameters are then used to determine the hidden states distribution in the next E-step iteration. The algorithm converges once the absolute value of the difference between parameters estimated in two iterations are less than some small amount[6].

4 Estimation

4.1. Data

Our data is sampled from one of the U.S. largest credit card issuers and includes data on credit card accounts and transactions. The transaction data contains credit card spending of each account from 2001 to 2003 and for each transaction it provides information on transaction date and amount as well as transaction category based on MCC. Account data contains credit card account-related information such as demographic characteristics of the account holder, monthly balance, APR, and repayment status for 9000 accounts. To speed up the runtime, we randomly sampled 2000 accounts from the data and after removing inactive accounts and accounts for which transaction data is missing or entirely zero, we are left with 1686 accounts for estimation. In order to avoid over-sampling from non-delinquent accounts and maintain the proportion of delinquency occurrences, we sampled separately from accounts with and without delinquency experience.

For each account the data provides information on the balance repayment status over a period of 25 months from July 2001 to August 2003. Account repayment status can be one of the following: (1) fully-paid the balance, (2) paid only the minimum payment, (3) paid some amount larger than minimum payment but less than the entire balance, and (4) delinquent. As mentioned in previous section, an account status is delinquent if a cardholder misses payments for two (or more) consecutive months. We define delinquency as a binary variable which takes the value of 1 if the account status is delinquent (and 0 otherwise) and use it as the response variable in our model. According to Netzer et al. (2008), the necessary condition for identifying

[6] In our model the tolerance is 10E-10

a hidden Markov model is the existence of dynamics in the response variable over time. To test the randomness of individual-level delinquency incidents, we used the Runs Test. The result of the test strongly supports the dynamics in delinquencies. In our data, account status is delinquent 13% of times overall accounts and time periods. Average utilization rate is 33% and the average credit card fee and monthly finance charge are $8 and $38, respectively. The average age of account holders is 43 years old with 13-year standard deviation which represents a sample of mostly 30-year old plus credit cardholders. 58% of the account holders have high-income jobs and about 23% of the cards are reward cards (see Table 1 for descriptive statistics of the data).

4.2. Variable Selection

Spending Behavior Covariates: include a set of variables capturing the monthly aggregate amount of spending in each of the transaction categories (MCC) denoted by vector (**W**) in our model. Due to large number of transaction categories (~ 300 different MCC) and limited number of observations for each category, we summarize MCC categories into larger groups based on the similarity of product/service being purchased. For instance, all different categories related to buying from discount stores, rental and used stores, or pawnshops are grouped under "Discount Shopping" category. By summarizing MCC we reduce the number of transaction categories from more than 300 to 28. However, the summarized categories still introduce many parameters to our model and cannot be used directly. Therefore, we choose from categories mainly based on two criteria. First, we choose the categories which have non-zero values for at least 5% of the times as due to the limited number of accounts in our data only 12 (out of 28) categories have non-zero values more than 5% of the times. Second, we use step-wise logistic regression to find the categories that improve the fit the most. We incorporate that with the discussions of purchase categories with highest correlation with credit risk in section 2, we narrow down to the 5 categories of discount shopping, entertainment, fees and payments, necessity goods, and travel and camping. Discount shopping include spending on discount stores, used merchandise and second-hand stores, variety stores, and wrecking and salvage yards. Entertainment category includes sub-categories such as betting and casino gambling, video games, bowling alleys, dating and escorting services, billiard and pool establishment. Fees and payments category consists of items such as automatic/manual cash disburse, utilities, and fines and governmental fees. Necessity includes spending gas, grocery, service stations, bakery and dairy products, and nondurable goods. Travel and camping contains any spending on hotels and accommodation, camping equipment, car rental, etc.

Table 1 - Summary Statistics

Variables	Mean	SD
Utilization Rate	0.332	0.411
Prior delinquency experience	0.120	0.173
Recency of delinquency	0.334	0.369
Fee	7.507	17.421
FC	37.989	60.754
job	0.588	0.492
age	43.342	13.471
rewards	0.225	0.418
Discount Shopping Spending	109.242	144.240
Entertainment Spending	78.883	137.569
Fees and Payments Spending	464.362	581.312
Necessity Goods Spending	146.808	198.966
Travel and Camping Spending	348.431	595.957
Percentage of delinquency	13%	
Number of Accounts	1686	
Number of Months	25	

Debt Repayment Covariates: include variables that describe the balance repayment behavior of customer in previous time periods denoted in the model by vector (**X**). Debt repayment behavior covariates are derived from the account data and include variables such as utilization rate, prior delinquency experiences, and recency of last delinquency incident. As suggested by Zhao et. al (2009), we define all the covariates in ratios to avoid any possible heteroscedasticity. Utilization rate at each time period t is defined as the ratio of balance over credit limit. Prior delinquency experience is the ratio of the total number of times a customer was delinquent divided by total number of time periods until $t - 1$. Similarly, recency is the number of time periods since last delinquency divided by the total number of time periods until $t - 1$.

Customer-Specific Characteristics: capture the observed heterogeneity across individuals and denoted by $\mathbf{Z}$ in our model. These variables are known at time 0 and do not change over time and mostly include account-specific or demographic factors. The variables we used are account holder age and job where job is defined as a dummy variable which is 1 if customer has a high-paying job and 0 otherwise.

4.3. *Estimation Results*

We find the maximum of the likelihood function using the Expectation-Maximization algorithm that also allows for the calculation of standard errors of parameter estimates. In order to estimate the hidden Markov model, the number of hidden states should be determined. Since we don't have any *a priori* information about the number of hidden states, we estimate the model for different number of hidden states and use AIC to find the number of states which provides the best fit (lowest AIC) to the data. AIC is defined as follows

$$AIC = -2\widehat{LL} + 2n$$

where $\widehat{LL}$ denotes the maximum of the log-likelihood function and n is the number of estimated parameters. In addition to hidden Markov model, we also estimate a simpler model to see if considering the hidden states would actually improve the fitness[7]. Specifically, we use a logistic regression model with the same response and explanatory variables as a benchmark to our hidden Markov model. Table 2 shows that a hidden Markov model with three hidden states provides the best fit to the data.

---

[7] In addition to improving model fit, we use hidden Markov model to use hidden states as a basis for risk evolution-based clustering of customers.

Table 2 - Model Fit and Prediction Measures

Model	AIC
Logistic model	27415
HMM – 2 states	27100
HMM – 3 states	26347
HMM – 4 states	29237

EM algorithm provides the estimated probabilities of being in each of the hidden states at each time period, which can be denoted as $p(\boldsymbol{S}_i^{(t)} = s_k | \boldsymbol{Y}_i^{(t)} = \bar{y}, \widetilde{\boldsymbol{Q}}_i^{(t)} = \bar{q})$, where k is the hidden state index. This can be particularly useful if we could determine the estimated state of customer and link it to the actual delinquency incident. In order to do so, we use a method called "global decoding" based on the algorithm suggested by Viterbi (1967). Viterbi method is a dynamic programming algorithm to find the hidden state that maximizes the probability of observed data. Table 3 shows the average of observed delinquency rate, $\boldsymbol{Y}_i^{(t)}$, overall customers, $i = \{1, \ldots, N\}$, and all time periods, $t = \{1, \ldots, T\}$, categorized by hidden states. In the next section, we further discuss the application of predicted states in explaining spending behavior variation across different risk profiles.

Table 3 shows that the rate of delinquency is 1.2%, 18.5%, and 81.4% for hidden states 1, 2, and 3, respectively. Therefore, we characterize hidden states as "no risk", "low-risk", and "high-risk" groups of observations where risk is defined as the probability of delinquency. It should be noted that the grouping is over hidden states of customers and not the customers themselves. This means that a customer can transition between different risk states throughout time based on his/her previous months' balance repayment behavior and current month's spending.

Table 3 - Conditional Posterior Credit Risk Probability

	Risk state = no risk	Risk state = low risk	Risk state = high risk
Delinquent = NO	0.988	0.815	0.186
Delinquent = YES	0.012	0.185	0.814

We also calculate the average posterior probability of transitioning between different hidden states (Table 4). The diagonal elements of the transition probability matrix have larger values implying varying degrees of stickiness in different risk levels. The stickiness is decreasing with the level of risk associated with each state, as lower risk profiles have higher tendency to remain in the same state than higher risk states. This suggests that most of the customers tend to evolve into a stationary lower risk state eventually, despite the initial random assignment to different risk states in our model. In terms of off-diagonal elements and the transitioning between different risk states, there is a large difference between upper and lower triangle as the average probability of moving from a lower risk state to a higher one is 0.03 versus 0.27 for moving from a higher risk state to a lower risk one. It shows that on average it's 9 times more likely to transition to a lower risk states than into a higher risk one. Moreover, Table 4 shows that probability of moving from the "no risk" to "high risk" state is almost zero whereas the probability of the reverse transition is 0.43. It suggests that the transition to "high risk" is not sudden (in one-time period) but the path from "no risk" to "low risk" and then to "high risk" is much more likely. In another words, becoming risky is more a function of a multi-period behavior rather than a one-time incident as it is also not sensible to profile a customer as risky for the first delinquency incident. Moreover, the higher probability of transitioning from "high risk" to "no risk", 0.43, than from "low risk" and then to "no risk", 0.10, suggests that delinquency occurrence is 4 times more likely to transition a customer into the no risk state.

Table 4 - Average Transition Probability Matrix

		State at *t*		
		No risk	Low risk	High risk
State at *t-1*	No risk	0.940	0.058	0.002
	Low risk	0.104	0.856	0.041
	High risk	0.437	0.256	0.307

Tables 5 shows the parameter estimates of our hidden Markov model. The parameter estimates show that higher credit card utilization rate increases the chance of delinquency but the odds of delinquency are higher when customer is already in “high risk” state. Similarly, prior delinquency experience and the recency of delinquency experience are associated with higher risk of delinquency in the future, where odds of delinquency are higher for customers in high-risk state. Credit card fees increases the probability of delinquency for both states almost equally, however, finance charges only increase the chance of delinquency for customers in the high-risk state. In terms of demographic characteristics, older people have lower probability of delinquency regardless of their current risk state. Job does not have any significant impact on delinquency risk in our estimation results. Having a reward card is associated with lower risk for customers in low-risk state.

Table 5 - Estimation Results

Variables	**Parameter Estimates**	
	State 2 (low risk)	**State 3 (high risk)**
Intercept (State 1)	**-3.692***	**-9.203**
Intercept (State 2)	**-2.040**	**-4.261**
Intercept (State 3)	**1.211**	0.477
Repayment Behavior until $t-1$		
Utilization Rate	**1.788**	**2.730**
Prior Delinquency Experience	**2.403**	**4.302**
Recency	**1.104**	**3.338**
Credit Card Fees	**0.020**	**0.026**
Finance Charges	0.000	**0.004**
Customer Demographics		
Job (high income = 1)	0. 118	-0.186
age	**-0.008**	**-0.009**
Reward Card (Yes = 1)	**-0.218**	0.018
Purchase Categories at time t		
Log(Discount Shopping)	**0.162**	-0.433
Log(Excessive Entertainment)**	0.438	**1.322**
Log(Fees and Payments)	**0.514**	0.012
Log(Necessity Goods)	0.053	**-0.696**
Log(Travel and Camping)	**0.175**	-0.262

* Bold fonts indicate significance at 5% level.
** Excessive means only purchases greater than mean+SD for that category is considered.

In terms of the impact of purchase category on delinquency risk, the results are diverse. For customers in low-risk state, spending in discount stores, pawnshops, rental and second-hand stores, home utilities, manual and automated cash disburse, travel and camping is associated with more risk. However, for high-risk customers, excessively[8] spending casinos and gambling, bowling alleys and billiard pools, dating services and game videos is associated with high delinquency risk. This is in line with findings of congress report on credit card profiling mentioned in section 2. Using credit card on necessity goods such as grocery, gas, and service stations is associated with lower risk.

4.4. Spending Behavior and Credit Risk Profile

We can predict the point estimate of hidden states sequence $s_i = \{s_1, \dots, s_T\}, s_t \in \{1, \dots, M\}$ for each individual $i = 1, \dots, N$ using maximum a posteriori (MAP) method. Based on the observed data, the MAP criterion finds a state s which maximizes the following,

$$\tilde{\boldsymbol{S}}^{*} = \underset{s}{\operatorname{argmax}}\, P\big(\widetilde{\boldsymbol{Q}}, \boldsymbol{Y} \big| \boldsymbol{S}\big) P(\boldsymbol{S}) \tag{6}$$

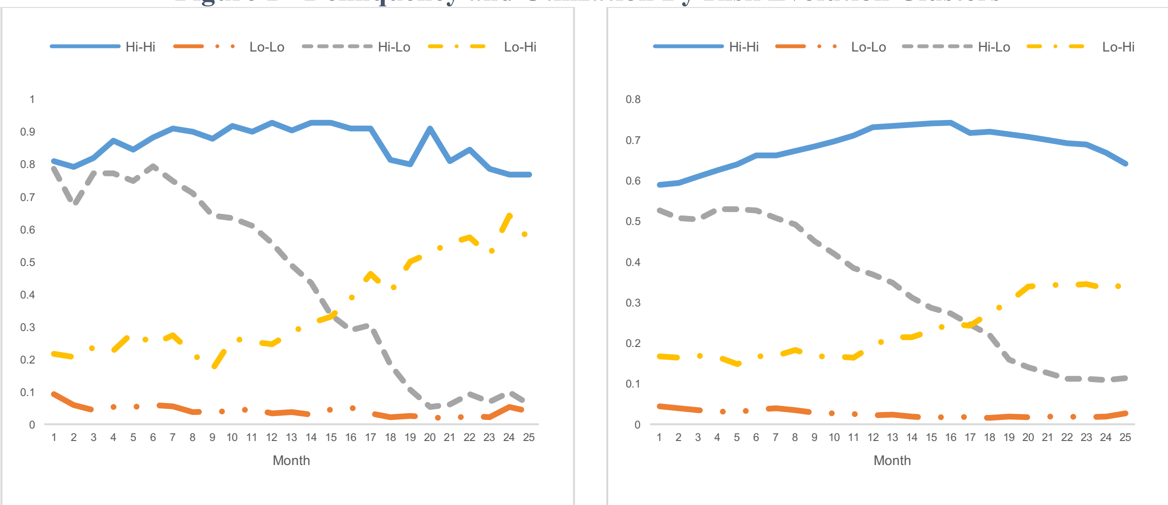
Where $\widetilde{\boldsymbol{Q}}, \boldsymbol{Y}$ are the observed data. The optimization criterion of MAP is similar to that of maximum likelihood method but MAP augments the objective function by incorporating the prior distribution of hidden state parameters. To find the solution to the optimization problem, we use Viterbi (1967) algorithm which maximizes the objective function without performing an exhaustive search of all M^Tpossibilities for each individual. The MAP estimation gives the matrix $\tilde{\boldsymbol{S}}$ of size $(N \times T)$ which shows how the risk state of each individual evolves through time.

We implement a cluster analysis on the predicted state matrix, $\tilde{\boldsymbol{S}}$, to find groups of customers with similar risk evolution pattern. K-means algorithm with 4 clusters explains about 77% of variance in the data where larger number of clusters improves the variance explained only marginally. The size of the risk evolution-based clusters is as follows: cluster1 459,

[8] Excessive spending in a category for an individual is defined as transaction values greater than mean plus standard deviation in that category.

cluster2 121, cluster3 897 and custer4 209. Figure 1 shows the evolution of the average of delinquency and utilization rate throughout the 25-month period of data for each cluster.

Figure 1 - Delinquency and Utilization By Risk Evolution Clusters



(a) Ave. delinquency incident (delinquent = 1) **(b) Ave. utilization rate**

Based on Figure 1, cluster1 represents the group of customers who are in high risk state from the beginning and continue to stay in it. We call this group "Hi-Hi". Cluster2, which is the largest cluster, is always in low risk state and we call them "Lo-Lo". Similarly, cluster3 represents the group of customers who start at high risk but lower their risk through time and cluster4 contains the group of customers who starts at low-risk states but end up at high risk one. We call them "Hi-Lo" and "Lo-Hi", respectively. As shown in Table 6, although Hi-Hi and Lo-Lo groups are the two highest and lowest risky customers, Hi-Lo and Lo-Hi customers show larger volatility in their credit risk evolution as they have much larger standard deviation.

Table 6 - Average Delinquency and Utilization Rate By Clusters

	Hi-Hi	Lo-Lo	Hi-Lo	Lo-Hi
Delinquency Rate	86% (5.5%)*	4% (1%)	44% (28%)	36% (14%)
Utilization rate	68% (4.5%)	2.5% (~ 0%)	33% (16%)	23% (7%)
Finance Charges	$64.6 ($19.7)	$5.86 ($2.64)	$60.96 ($12.64)	$18.25 ($3.75)
Credit Card Fees	$12.4 ($2.18)	$1.33 ($0.48)	$9.98 ($3.21)	$4.47 ($1.19)

* numbers in parenthesis are standard deviations.

In terms of profitability, Hi-Hi and Hi-Lo customers have the highest finance charges and fees, however, Hi-Hi customers' high delinquency rate makes them unfavorable. Lo-Lo customers are the least risky and least profitable customers as they incur lowest finance charge and fees.

Table 7 - Transaction Category by Risk Clusters[9]

Category Name	Hi-Hi	Lo-Lo	Hi-Lo	Lo-Hi
Airlines	12%	37%	21%	30%
Car purchase	10%	**47%**	16%	27%
Car repair	12%	29%	20%	38%
Car parts	15%	42%	11%	33%
Books and papers	12%	21%	11%	**56%**
Clothing	16%	21%	16%	47%
Computer & electronics	9%	35%	16%	40%
Consulting services	**24%**	**49%**	6%	21%
Department store	15%	26%	19%	40%
Direct marketing	19%	26%	18%	37%
Discount shopping	14%	27%	**25%**	35%
Drinking	12%	30%	15%	43%
Durable goods	0%	**69%**	0%	31%
Educational expenses	13%	19%	16%	**51%**
Entertainment	**22%**	35%	13%	31%
Fees and payments	**62%**[*]	10%	12%	16%
Financial services	0%	**73%**	0%	27%
General services	13%	24%	**22%**	42%
Health expenses	12%	27%	11%	50%
Home fixing	10%	**73%**	0%	17%
Home improvement	11%	36%	**23%**	31%
Luxury goods	**20%**	32%	21%	27%
Membership costs	8%	31%	6%	**55%**
Music	16%	17%	11%	**55%**
Necessity goods	16%	39%	14%	31%
Personal care	**21%**	27%	17%	35%
Personal finance	8%	26%	17%	49%
Postal services	0%	31%	0%	**69%**
Restaurants and eating	14%	27%	16%	43%
Specialty goods	6%	41%	10%	43%
Sports	10%	24%	**32%**	34%
Telecommunication	18%	30%	14%	38%
Transportation	15%	35%	**24%**	26%
Travel & camping	9%	32%	16%	42%

[9] Top 5 categories in each cluster is bolded.

We can also look at the amount of spending in each transaction category across the 4 clusters. Table 7 shows the share of median spending in each category across all clusters[10]. The results show that customers in Hi-Hi cluster use their credit card for manual or automated payments of bills and utilities, and government fines (fees and payments) much more than other clusters. Customers in Lo-Lo cluster spends much higher in getting financial services, fixing home-related issues, and buying durable goods (mostly household appliances). Home fixing includes services such as renewing floor covering, repairing heating, plumbing, and AC issues, etc. Hi-Lo group spends mostly in sports, discount shopping as well as transportation. Lo-Hi cluster share of spending is around 50% and higher in several categories including music, membership fees, educational expenses, cloth shopping, drinking and eating at restaurants/bars. Lo-Hi cluster also has the largest amount of spending among other cluster whereas Hi-Hi and Hi-Lo average spending is the lowest.

4.5. Managerial Implications

In order to reduce the charge off costs, bank started mining purchasing patterns of customers in the hope of finding purchasing patterns that indicate imminent future risk elevation. Identification of purchase categories associated with high chance of delinquency helps banks to act proactively adjust APR or credit limit to minimize their cost if a charge off occurs. Our methodology helps to dynamically classify customers into risk groups and profile them based on their purchasing pattern. In addition to that, our model enables banks to identify the most and least profitable segments of customers. Based on our results, most profitable customers are the ones who revolve balance through time but don't default. Hi-Lo cluster fits in this category as they pay the highest finance charges but have a moderate delinquency rate

[10] median is used to avoid the effect of outliers.

(33%). On the other hand, Hi-Hi customers are the riskiest cluster for the bank despite their high finance charges because their delinquency rate is very high (86%). In addition to that, profiling each cluster based on their spending behavior directly links credit risk to purchase category which can be used as another way of identifying risk level change in a customer. For instance, if a customer has been spending relatively higher in the discount and pawn shops for some time periods, bank should assign a future risk flag to the customer's account and if necessary adjust APR or credit limit accordingly.

5 Conclusion

In this paper we look at the relationship between spending behavior and credit risk. We are interested to see if certain purchase categories are associated with higher credit risk and whether there is variation amongst purchase categories based on different levels of risk. We develop a hidden Markov model to estimate the impact of debt repayment and spending behavior on the probability of delinquency. We also use the estimated hidden states for a cluster analysis to find groups of people with similar risk evolution patterns.

We find that impact of different repayment covariates (e.g. utilization rate) on credit risk are the same, however, the magnitude of impact differs based on the current hidden risk state. We identified four clusters of customer based on their risk state evolution through time. Interestingly, each cluster showed different characteristics in terms of risk evolution and potential profitability for the bank. We find that the cluster of customers who start with high utilization rate and high risk states but evolve into lower risk states are the most profitable for the bank because they pay the highest finance charges while having a modest, decreasing level of risk. Also, we find that there is a segment of customers who starts with high utilization and delinquency rate and stay in the same state. They carry large balances and have rather high likelihood of default and should be the first priority of bank's risk management actions. We also show that highest spending categories are totally different among different risk clusters which indicates a relationship between customer credit risk and spending categories.

Our research has some limitations mainly due to lack or shortage of data. We haven't accounted for the share of wallet as we only have access to credit card transaction of customers from one bank. Having data on customers' spending with their other cards would provide more comprehensive view of their spending patterns as some customers may push some specific purchases to their other cards. Also, we didn't have enough observations in each MCC transaction category and were not able to see the direct impact of spending categories on delinquency.

References

[1] Basanik, J., Crook, J., Thomas, L.C. (1999), Not if but When will Borrowers Default?, *The Journal of Operational Research Society*, 50 (12), 1185-1190.
[2] Baum, L.E., Petrie, T., Soules, G., Weiss, N. (1970), A Maximization Technique Occurring in the Statistical Analysis of Probabilistic Functions of Markov Chains, *The Annual of Mathematical Statistics*, 41 (1), 164-171
[3] Bellotti, T., Crook, J. (2009), Support Vector Machines for Credit Scoring and Discovery of Significant Features, *Expert Systems with Applications*, 36 (2), 3302-3308
[4] Danziger, P. (2004), Why People Buy Things They Don't Need, Paramount Market Publishing
[5] Dempster, A.P., Laird, N.M., Rubin, D.B. (1977), Maximum Likelihood from Incomplete Data via the EM Algorithm, *Journal of the Royal Statistical Society,* 39 (1), 1-38
[6] Hsieh, N. (2004), An Integrated Data Mining and Behavioral Scoring Model for Analyzing Bank Customers, *Expert Systems with Applications*, 27, 623-633
[7] Im, J.K., Aply, D.W., Qi, C., Shan, X. (2012), A Time-dependent Proportional Hazards Survival Model for Credit Risk Analysis, *Journal of Operational Research Society*, 63, 306-321
[8] Khandani, A.E., Kim, A.J., Lo, A.W. (2010), Consumer Credit Risk Models via Machine Learning Algorithms, *Journal of Banking & Finance,* 34 (11), 2767-2787
[9] Leemis, L.M., Schimeiser, B.W., Evans, D.L. (2000), Survival Distributions Satisfying Benford's Law, *The American Statistician*, 54 (3), 1-6
[10] Lin, X., Zhong, Y. (2012), An Overview of Personal Credit Scoring: Techniques and Future Work, *International Journal of Intelligence Science*, 2 (4), 181-189
[11] Netzer, O., Lattin, J.M., Srinivasan, V. (2008), A Hidden Markov Model of Customer Relationship Dynamics, *Marketing Science*, 27 (2), 185-204
[12] Pacelli, V., Azzollini, M. (2011), An Artificial Neural Network Approach for Credit Risk Management, *Journal of Intelligent Learning Systems and Applications,* 3 (2), 103-112
[13] Pavlenko, T., Chernyak, O. (2010), Credit Risk Modeling Using Bayesian Networks, *International Journal of Intelligent Systems*, 25 (4), 326-344
[14] So, M.M.C., Thomas, L.C. (2011), Modelling the Profitability of Credit Cards by Markov Decision Process, *European Journal of Operational Research,* 212 (1), 123-130
[15] Stepanova, M., Thomas, L.C. (2002), Survival Analysis Methods for Personal Loan Data, *Operations Research,* 50 (2), 277-289
[16] Tang, D.Y., Yan, H. (2007), Liquidity and Credit Default Swaps Spreads, *Available at: SSRN: http://ssrn.com/abstract=1008325*
[17] Thomas, L.C. (2000), A Survey of Credit and Behavioral Scoring: Forecasting Financial Risk of Lending to Consumers, *International Journal of Forecasting*, 16 (2), 149-172
[18] Thomas, L.C., Oliver, R.W., Hand, D.J. (2005), A Survey of the Issues in Consumer Credit Modelling Research, *Journal of Operational Research Society,* 56 (9), 1006-1015
[19] Thomas, L.C., Stepanova, M. (2001), PHAB Scores: Proportional Hazards Analysis Behavioral Scores, *Journal of Operational Research Society,* 52 (9), 1007-1016
[20] Vissing-Jorgensen, A. (2011), Consumer Credit: Learning Your Customer's Default Risk from What (S)he Buys, *working paper available at SSRN: http://ssrn.com/abstract=2023238*
[21] Zhao, Y., Zhao, Y., Song, I. (2009), Predicting New Customers' Risk Type in the Credit Card Market, *Journal of Marketing Research*, XLVI, 506-517